Commitment Contracts and School Competition for Smoking Prevention in Indonesia*

Margaret Triyana† Justin S. White‡

October 31, 2018

Abstract

We provide evidence on the effectiveness of a non-monetary incentive program to promote healthy behavior among adolescents. This school-based field experiment invited students and parents to sign a pledge for students to abstain from smoking and for parents to monitor their children. The program included a non-monetary penalty and regular monitoring in schools, and a subset of participating schools competed against each other for the highest tobacco abstinence rates as a means to test team incentives that harness peer effects. We find that the individual pledges increase biochemically verified tobacco abstinence by 5 percentage points, an effect that is sustained three months after the end of the intervention. School competition has no additional impact on tobacco abstinence. Our findings highlight the effectiveness of non-monetary incentives to curb unhealthy behaviors among adolescents who face problems with self-control and peer pressure.

Keywords: Smoking prevention, commitment contracts, incentive-based program, school-based intervention, health behavior, peer effects

JEL codes: C93, D91, I12, O10

*We thank Dr. Hasto Wardoyo, Baning Rahayujati, and 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 seminar participants at the University of Notre Dame, the Population Association of America meeting, Luksburg PUC-ND Conference, and Population Health Sciences Research Workshop at Boston University. We acknowledge funding from J-PAL Southeast Asia, 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.

†University of Notre Dame

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

1 Introduction

Encouraging healthy behavior has the potential to improve long-term well-being by preventing the development of serious health conditions. Yet, in both high- and low-income settings, people tend to underinvest in preventive health behaviors [Kenkel, 1994, Dupas, 2011]. To increase health investments, incentive programs that employ rewards and penalties have become increasingly common in the health domain, including in low- and middle-income countries (LMICs) [Gopalan et al., 2014]. In this paper, we test the effectiveness of a school-based incentive program to prevent smoking in early adolescence in Indonesia.

Tobacco use is associated with 6.4 million annual deaths worldwide, with nearly three-quarters occurring in LMICs [Reitsma et al., 2017]. The high disease burden associated with tobacco use has led to various strategies such as increasing taxes on tobacco products [World Health Organization, 2015c]. Price, however, has not been a politically feasible policy lever in Indonesia, a tobacco-producing country where 70% of men smoke and about 15% of deaths are attributed to tobacco use [World Health Organization, 2012c]. Due to these policy constraints, policymakers in Indonesia are increasingly interested in implementing non-price tobacco control strategies, including prevention programs. Our study implemented a non-monetary commitment contract for smoking prevention among adolescents.

Incentive programs have been shown to be effective in changing a number of health behaviors.[1] To boost the effectiveness of incentives, several studies have tested commitment contracts for individuals to voluntarily make tempting options more costly in the future, often by putting their own money at risk, conditional on achieving a health goal [Bryan et al., 2010, Giné et al., 2010, Halpern et al., 2015, Royer et al., 2015, Bai et al., 2017, Schilbach, Forthcoming].[2] The commitment literature has focused almost exclusively on adults, providing little guidance on the effectiveness of commitment mechanisms among adolescents.[3] 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]. Consequently, adolescents are susceptible to engaging in risky behavior such as smoking. A better understanding of the impacts of commitment mechanisms among adolescents may improve policy design and implementation,

[1] See Mantzari et al. [2015], Giles et al. [2014], and Gopalan et al. [2014] for recent reviews of the literature.

[2] 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.

[3] A few studies outside of the health domain have targeted formal commitment devices to adolescents [Ariely and Wertenbroch, 2002, Acland and Chow, 2018].

with an opportunity to intervene at a critical stage in the formation of habits and preferences.

In this study, we test whether a non-monetary commitment mechanism effectively prevents smoking among adolescents in Indonesia, and we explore potential mechanisms associated with the program effects. Our intervention also leveraged adolescents' various social networks. The intervention engaged teachers (school), parents (family), and student peers (friends) to test the effectiveness of commitment, parental influence, and peer effects. Specifically, we implemented a school-based field experiment that involved about 2,700 13-year old students from 72 schools in Indonesia. In the treated sample of 36 schools, students in each school were invited to publicly sign a pledge in which they agreed not to smoke, and their parents were invited to privately sign a pledge to monitor their children for tobacco use at home. Many Indonesian schools employ a disciplinary system wherein teachers issue demerit points to students who misbehave. We took advantage of this system by stipulating in the commitment agreement that students caught smoking would receive demerit points at school and a warning from the guidance counselor for the first two infractions. In addition, the third and subsequent infractions would result in parental notification, which schools believed would constitute a sufficiently severe penalty for tobacco use.[4] Throughout the 5-month intervention, schools collected bi-weekly reports from students in 7th grade (and their teachers) to identify smokers and remind students of the importance of avoiding tobacco use. In 18 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 schools received a bi-weekly message that congratulated them if their tobacco abstinence rate was in the top ten of participating schools. At the end of 5 months, the local government officially recognized the top ten schools with a plaque.

We conducted an initial follow-up at 3 months to evaluate the immediate program effects. Using an intent-to-treat approach, we find an approximately 5-percentage point increase in biochemically verified smoking abstinence at 3 months. We conducted a second follow-up at 8 months (about 3 months after schools had ceased bi-weekly monitoring) to evaluate the medium-term program effects. We find that the effect persisted to 8 months, when intensive monitoring was no longer in place. These results suggest that the bi-weekly monitoring and reminders are not the sole cause of the program effect. Overall, these results provide evidence for the medium-term effectiveness of a non-monetary commitment device in promoting healthy behavior among adolescents.

[4]Before our intervention, tobacco use would typically result in a warning from the guidance counselor, and in schools with pre-existing demerit point systems, the demerit points were generally lower than other disciplinary infractions.

With regard to the school competition, our analysis reveals that the school competition had no additional effect on smoking abstinence rates. The individual commitment and school competition interventions produce similar intent-to-treat effects after 3 months. At 8 months (3 months after bi-weekly monitoring had ceased), the effects of the individual commitment and school competition remain similar in magnitude. The school competition tests whether social effects, including emphasizing group identity, changes behavior more effectively than the individual pledge alone. We fail to find evidence in support of this hypothesis since the individual commitment, rather than the school competition, drives the program effect.

Given the high smoking rates in many LMICs and the addictive nature of tobacco use, a scaled-up version of our program would have the potential to have a large impact on tobacco prevalence and the tobacco-related disease burden. Our study contributes to a growing literature in behavioral and health economics on the use of incentives, commitment contracts, and peer effects to improve health outcomes such as weight loss, diabetes management, and smoking cessation. We depart from the existing commitment literature to provide evidence for the effectiveness of the use of an incentive mechanism among adolescents whose executive function is still developing and who are susceptible to risky behavior such as smoking. While prior studies have encouraged students to make a personal promise not to smoke [Sussman et al., 2007], our commitment contract is different in three important ways. First, our program employs a non-monetary penalty. In the behavioral economics literature, commitment contracts have typically been structured to put an individual's own money at stake. In our study, however, we adopt a non-monetary penalty (demerit points and parental notification) if students are caught smoking, thereby improving the chances that the intervention is sustainable, culturally acceptable, and scalable.

Second, individual participants are often solely responsible for enforcing and remaining committed to their individual pledge. We, on the other hand, engage a variety of individuals (parents, teachers, and fellow students) to monitor and enforce the individual pledge to avoid smoking, a risky behavior. Risky behavior in adolescence is 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 in which they observe their parents and peers engage [Bandura, 1986]. Since parents and peers play an important role in influencing adolescents, they also have the potential to play a key role in mediating adolescent risky behavior, particularly tobacco use [Lundborg, 2006, Trogdon et al., 2008, Ennett et al., 2001, Kremer and Levy, 2008, Card and Giuliano, 2013].

Third, we implement the smoking 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, all but four of the 49 studies included in the analysis took place in high-income countries[5]. 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. Section 4 describes the methods used to estimate the intervention's effects. Section 5 presents our findings. Section 6 concludes with policy implications.

2 Experimental design

We implemented a cluster randomized control trial where schools served as the unit of randomization in order to understand the effects of non-monetary incentives and school competition on smoking behavior among middle-school children in Indonesia. We used schools as the unit of randomization to minimize the chance of spillover effects among classmates and to increase perceived acceptability by school officials. Our pre-analysis plan is available on the AEA RCT Registry (AEARCTR-0001607).

2.1 Setting

In Indonesia, tobacco use is a particularly important public health issue because it accounts for 15% of all deaths [World Health Organization, 2012a], specifically 20% of deaths among men and 12% of deaths among women (mostly from second-hand smoke) [World Health Organization, 2012b]. In addition, Indonesia has the highest male smoking prevalence in the world at about 70% [World Health Organization, 2012c, 2015a]. In fact, nearly 12% of total expenditures in households with a smoker go to tobacco products [World Health Organization, 2012c, 2015a]. In spite of the high burden of disease, Indonesian smokers have shown little interest in smoking cessation [Nichter et al., 2009, 2010], which indicates that alternative tobacco control strategies are necessary to reduce the tobacco-attributable disease burden.

The government's reliance on the tobacco industry as a source of government revenue and employment has largely contributed to Indonesia's tobacco problem. The tobacco industry

[5] Two studies were in China, one in South Africa, and one in Thailand.

accounts for about 10% of total tax revenue in Indonesia[6] and is the second largest employer in the country [Barber et al., 2008]. As such, the industry wields significant political and financial power. Consequently, price has not been a feasible policy lever. Non-price tobacco-control policies are also limited in Indonesia, as evidenced by Indonesia's status as one of the few nations that has not 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 [Ng et al., 2007b]. 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, 2012c, Lillard et al., 2013]. In Indonesia, the trend is similar: adult smokers initiated smoking at age 17 on average, but age of initiation has been decreasing [World Health Organization, 2012c]. According to recent national data, adolescents report first smoking between the ages of 12 and 13 [World Health Organization, 2015b]. 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 how adolescence is a critical period for smoking prevention in Indonesia.

In Indonesia, the wide-ranging health risks associated with smoking tend to be downplayed. Among the Indonesian population, it is widely believed that smoking 5-10 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 5-10 cigarettes per day is not particularly harmful for health reinforces this mentality [Ng et al., 2007a]. 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. Therefore, Indonesia's culturally high smoking rates and policy environment provide an ideal setting to test innovative tobacco control strategies.

Our study was implemented in two districts in Yogyakarta Province in the Central Java region: Kulon Progo and Sleman. Kulon Progo has adopted some of the most restrictive

[6]https://nasional.kompas.com/read/2018/03/20/23224701/penerimaan-cukai-tembakau-terus-meningkat. Last accessed September 27, 2018.

tobacco control policies in Indonesia, while Sleman's are less restrictive due in part to the fact that some sub-districts in Sleman produce tobacco.[7] 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, as well as tobacco advertising and event sponsorship from tobacco firms. At the start of this study, we partnered with local government agencies in each district to help them implement their tobacco control policies in schools. Our intervention design was informed by extensive consultations with the local governments to ensure its acceptability and sustainability.

Smoking among females is rare in Indonesia, about 2% of women smoke.[8] Therefore, our study targeted 12- to 13-year-old male students in the 7th grade (*Sekolah Menegah Pertama*) who typically graduate at the age of 15-16 years by the end of 9th grade. In 2013, among smokers above the age of 10, 36% of smokers in Kulon Progo and 45% in Sleman began smoking between 15 and 19 years of age, while 13.4% and 9.2% began smoking between the ages of 10 and 14 years in Kulon Progo and Sleman, respectively.[9] In our study, most male students report smoking for the first time between the ages of 11 and 13 years, indicating that the 7th grade is a critical period for intervention.

Schools in our sample have, on average, four classes per grade, and each class has a maximum of 42 students. Most schools in Indonesia engage with parents four times per year: parent-teacher meetings are held twice per year, and parents come to schools to pick up their children's report cards at the end of each semester. All schools in our sample have already instituted some disciplinary action on smoking. If a student is caught smoking, then most schools require students to attend a counseling session in school.

2.2 Interventions

Figure 1 summarizes the study design. There are 36 schools in the control arm, 18 schools in the individual commitment arm, and 18 schools in the school competition arm. A total of 2,984 students were screened, and 256 were excluded. 1,315 students were allocated to the control arm, 668 students were allocated to the individual commitment arm, and 745 students were allocated to the school competition arm. All participating schools received an information campaign on the risks of smoking, comprised of one 45-minute session per class

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

[8] `http://www.who.int/tobacco/surveillance/policy/country_profile/idn.pdf`. Last accessed October 15, 2018

[9] Riset Kesehatan Dasar 2013.

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 preexisting 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 an agreement to abstain from smoking during the intervention period.[10] After students signed their individual commitment contracts, they were invited to sign an additional agreement 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 contract stating that the parents would monitor their children to prevent them from smoking. Students who violated the terms of the contract, as measured by a biochemically verified cotinine test and student and teacher reports, were penalized using a school-based demerit system.[11] Upon the first two infractions, the student was summoned to the guidance counselor's office. 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.

Guidance counselors, members of the student council, and classmates were all tasked with monitoring and reporting participating students' smoking behavior. Most schools in Indonesia have student councils, who are elected in each academic year. Students from the 7th, 8th, and 9th grades would elect nominated 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 or 9th grade), who are widely viewed with respect by other students. Moreover, due to their tendency to be academically minded, smoking rates tend to be very low among this group. The student council would gather reports from students and teachers regarding smoking behavior among children in participating classes. Every two weeks, students in the 7th grade were asked to complete a form indicating any students who they observed smoking on or off school premises. These bi-weekly reports served to bring smoking prevention to students' top of mind. The reports were kept private (except as part of an appeals process,[12]) but they were not strictly anonymous. The reporting mechanism, in which all students provided reports at the same time, was

[10] The student and parent contracts are provided in Appendix Figures B.1 and B.2.

[11] The demerit system for our study was kept separate from the school's pre-existing demerit system, in order to standardize the program. Our demerit system did not result in academic suspension or expulsion. All students in 7th grade can be reported regardless of whether they signed the contract. However, those who did not sign the contract would not receive the demerit points upon failing to comply.

[12] The reporter's identity may be revealed to the guidance counselor as part of an appeals process. A student who is accused of smoking could appeal to the guidance counselor, who would investigate. No appeals were requested throughout the study.

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 on the school premises. Students could also report cases directly to the student council, teachers, or our program administrators.

In the treatment group, members of the student council received training on how to promote smoking abstinence among participating students.[13] The student council, supervised by the guidance counselor, would receive, investigate, and record reports of students who smoked. These reports would ultimately be kept by the guidance counselor to be evaluated, and if necessary, be followed up using the demerit system.

The school competition was scored based on the bi-weekly reporting, the results of an audit during a randomly timed study visit to each school during the intervention period, and the biochemical test results at endline.[14] The school's score included several components: timeliness of reporting, completeness of reporting, and the proportion of students who were smoke-free.[15] We provided a list of the top 10 ranked schools on an ongoing basis. The schools that finished in the top 10 received a certificate of recognition from the district's Department of Education.

We had two key study hypotheses. First, commitment contracts would reduce tobacco use among students, compared to receiving information only. We expected that a combination of factors would curb smoking behavior, including monitoring from parents and teachers, enforcement of in-school remediation for smoking, and enhanced motivation and self-efficacy not to smoke due to students' commitment. Second, class competition combined with a commitment contract would reduce tobacco use among students, compared to a commitment contract alone or receiving information alone. The competition was expected to motivate students not to smoke, create peer pressure from classmates not to smoke, and engage

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

[14] One potential concern is schools in the competition arm would have an incentive to under-report. We find similar rates of reporting across the individual commitment arm and the school competition arm at 3 months and 8 months, when bi-weekly reporting had ended.

[15] We used the following formula to calculate bi-weekly school performance:

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

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

students to establish a norm against smoking.

2.3 Sample and time line

The intervention was implemented in 19 sub-districts in 2 districts in Yogyakarta province. The sub-districts were selected to be similar in their socio-economic characteristics. The study involved a total of 72 schools: 36 schools in Kulon Progo and 36 schools in Sleman. We used pair matching 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 conducted sample size calculations based on pairwise comparisons between equal-sized study arms. Based on previous smoking prevention studies, our study used a minimum detectable effect of 5 percentage points.[16]

Figure 2 describes the study timeline. The study started in January 2017, which corresponded to the second half of the academic year 2016-2017. We collected bi-weekly reports from all treated schools between January and May 2017. Reporting ended in May due to the beginning of the national examination period for all schools. We conducted an unannounced audit in all 72 schools in March 2017, 3 months after the intervention started. We conducted biochemical verification for about 80% of students. In the 36 treated schools, this audit also served to check the bi-weekly reports. 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, all students from that school received the audit. In larger schools, we randomly selected 48 students.[17] After the audit, bi-weekly reporting continued through May 2017.

The academic year ended in June 2017, and the new academic year began in July 2017. Students were randomly assigned new classmates when they entered 8th grade. We leverage this source of variation in our analysis of peer effects when the students entered 8th grade. Bi-weekly reporting could not resume at the beginning of the new academic year because the student council had not yet been formed for the year. Another round of cotinine testing was conducted in August 2017, about 8 months after the intervention had started. We use this follow-up to evaluate the longer-term program effect after monitoring had ceased.

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

[17] This was done to maintain an average of 30 participants per school for the unannounced audit.

3 Data

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. The baseline survey was conducted in January 2017. A spot check was conducted in March, approximately 3 months post intervention. The end line survey was conducted in August 2017, in the beginning of the new academic year.

Outcomes Our primary outcome is the proportion of students who abstained at 8 months. 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 urinary cotinine at concentrations $\geq$ 200 mg/ml. According to the package insert, the test is also highly sensitive at 100 mg/ml. The window of detection is expected to be up to 2 to 3 days after nicotine use. The urine tests were conducted at baseline, during the unannounced audit, and at endline. Second, students provided self-reports of their smoking status as part of each round of survey. Third, the student council, teachers, and fellow students were asked to report the names of students from the school who had smoked in the prior week during the random audit and at endline.

We combine the cotinine results and self-reports of smoking status to generate a combined smoking status variable, 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 abstinence is verified with the cotinine test. In other words, if a student reported abstinence, but failed the cotinine test, we code the student as failing to abstain. In cases where smoking status is missing, we drop observations with missing outcome data. The rate of missing cotinine test is similar across treatment and control at baseline and 8 months.[18] Similarly, the rate of students who self-reported abstinence but failed the cotinine test is similar across treatment and control at each wave of the survey.[19] We test for robustness by estimating models in which we assume students with missing outcome

[18]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.

[19]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 (o-value 0.60) at 8 months.

data are non-smokers unless they are observed or reported to have smoked at some point during the trial.

Additional outcomes include tobacco knowledge, the frequency of being reported as smoking, and smoking intentions. Their smoking intention is based on their prediction of the likelihood of smoking within 3 months, reported on a scale of 0 to 10. Tobacco knowledge is based on a score on a series of 9 questions to assess students' awareness of the health risks of tobacco use. Smoking intention 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. We also include truancy as a related disciplinary outcome.

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 school survey includes characteristics, such as: the number of teachers who smoke in the school, timing of the last parent-teacher conference, the presence of disciplinary action against smoking, and the average national exam scores in mathematics. 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 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 ownership of household assets. We also ask students to predict the likelihood of each of their five closest classmates to start smoking within 3 months, reported on a scale from 0 to 10. The student survey also includes time and risk preferences, where students were given hypothetical gambles to elicit their preference for certainty and time inconsistency. Students were also asked whether they believed their parents would find it acceptable if the student smoked to capture students' belief of parents' attitude towards smoking.

The parent survey includes household size, home ownership, land ownership, the occupation of the head of the household, the number of smokers in the household, and whether at least one parent smoked. The occupation of the head of the household is categorized into agriculture, self-employment in non-agriculture, government employee, blue collar work, and private employment in non-agriculture. We define skilled occupation as self-employed in non-agriculture, government employee, or privately employed in non-agriculture. We also create an indicator for low paternal (maternal) education when the child's father (mother) has primary education or lower, corresponding to less than 6

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

years of education. To capture parents' smoking attitude, we also include an indicator that takes the value one if parents find it unacceptable if the child smoked.

3.2 Sample statistics

Overall, baseline characteristics are similar across treatment and control groups (Table 1). Panel A describes the school characteristics and Panel B describes student characteristics at baseline. Schools in our sample are about 17 kilometers away from the main city of Yogyakarta. About 25% of schools in our sample are urban. The average number of students in the school is about 360. The average student to teacher ratio is about 0.073 (or 14 students per teacher) in both control and treatment groups. All the schools in our sample have student councils, comprising of about 10 students. We use the average mathematics score on the national examination as a proxy for school quality.

Most students abstained in the seven days prior to the baseline survey, implying that about 22% of students had smoked by the time they reached 7th grade. Most students who smoke choose clove cigarettes (kreteks), which are the most commonly found form of tobacco in Indonesia, followed by the e-cigarettes. More than 50% of students who reported smoking within the past month had smoked for the first time between the ages of 11 and 13. Based on the baseline biochemical verification, about a third of students smoked over the weekend, which suggests non-school influences may affect students' decision to smoke.

The average household size is about 5. About a third of the head of the household are in skilled occupations. Almost half of students have at least one parent who smoked, typically the father since smoking among women in Indonesia is rare. About half of households in the sample own land, which we use as our proxy for wealth. About 13% of mothers and fathers of students in our sample have primary education or lower.

4 Empirical strategy

4.1 Estimation of treatment effects

Our key empirical models test how the interventions affect tobacco abstinence. Our main outcome variable is verified smoking status, which combines the cotinine test and self-reported measure. We use a pooled analysis to assess the effects of the commitment contract using a binary treatment variable. We then use a sub-treatment model to assess the incremental effect of school competition. We report covariate-adjusted estimates.[21] The

[21] We also estimate unadjusted models, which only include the treatment indicator(s), with standard errors clustered at the school level. Results are reported in the Appendix (Table B.3).

multivariate analysis takes the following form.

Pooled analysis:

$$Y^*_{ist} = \alpha + \beta T_s + X_{is0}\gamma + \varepsilon_{ist}$$

Sub-treatment analysis:

$$Y^*_{ist} = \alpha + \beta_1 D_s + \beta_2 C_s + X_{is0}\gamma + \varepsilon_{ist}$$

where Y^*_{ist} takes the value one if person i attending school s at time t abstains from tobacco use. T_s is an indicator variable equal to 1 if the school is randomly allocated to a treatment group, D_s is an indicator variable equal to 1 if the school is randomly allocated to a commitment contract only (i.e., with no school competition), C_s is an indicator variable equal to 1 if the school is randomly allocated to a commitment contract plus school competition, X_{is0} is a vector of individual and school characteristics measured at baseline ($t = 0$), and ε_{ist} is the disturbance term. We include district fixed effects to capture non-time varying district characteristics.[22] Standard errors are clustered at the school level. The models above are estimated as logit specifications, and reported as average marginal effects. Non-binary outcomes are estimated using ordinary least squares. Standard errors are clustered at the school level to account for intra-cluster correlation of outcomes. We estimate the program effects at three and eight months after the beginning of the program.

Baseline covariates X_{is0} include student and school characteristics, the class average of the outcome variable, and baseline individual outcome. We calculated class average by calculating the mean after excluding the individual's outcome at baseline. Partially adjusted models include district fixed effects, the baseline class average, and individual outcome variable at baseline. Fully adjusted models add the following student characteristics: household land ownership, parental smoking status, an indicator for skilled occupation of the household head, an indicator for low maternal and paternal education, and the following school characteristics: the number of teachers who smoke and the average mathematics score on the national examination.

For robustness, we employ a difference-in-differences strategy to estimate the program effects by estimating the following equation:

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

where $Post_t$ is an indicator variable equal to 1 for the follow-up period, T_s is an indicator variable equal to 1 if the school is randomly allocated to a treatment group. Similarly, we

[22]Including pair fixed effects yields similar estimated effects.

conduct the following sub-treatment analysis using the following equation:

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

where D_s is an indicator variable equal to 1 if the school is randomly allocated to a commitment contract only (i.e., with no school competition), C_s is an indicator variable equal to 1 if the school is randomly allocated to a commitment contract plus school competition. All standard errors are clustered at the school level. The same covariates are included for partially adjusted and fully adjusted models.

4.2 Verification of experimental validity

There are two primary threats to the empirical design. First, the randomization may have produced 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. The primary independent variable is the intent-to-treat random allocation of schools to the treatment and control group, a dichotomous variable. A secondary trichotomous measure captures whether a school is allocated to the control arm, the commitment contract only arm, or the commitment contract plus competition arm. We test for balance along pre-intervention characteristics using a t-test of equality across arms and in bivariate comparisons that regress each covariate on a dummy equal to one if the intervention was implemented in the school. The differences between the treated and control groups are generally small and not statistically significant at baseline (Table 1). The average mathematics score among control schools is slightly higher, but the difference is not statistically significant. The joint test 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 not statistically significant. The joint test fails to reject the null that the sub-treatment arms are different from the control group.

Second, it is possible that students were not randomly re-assigned new classmates in 8th grade. We estimate an indicator that takes the value one if he is in the treated group and is assigned the same class in 7th and 8th grades (i.e., moving from 7A to 8A). Students who were assigned to different classes are on average similar to those assigned the same class at baseline, thereby allowing us to use their 8th grade peers as an additional source of variation (Table B.2). In addition, students who were assigned the same class and different classes have similar abstinence rates at 3 months after the beginning of the intervention.

5 Results

5.1 Effects on smoking abstinence

We examine the pooled program effect on verified abstinence in Figure 3 (Panel A). The unadjusted estimates are imprecise at three and eight months (Table B.3). Including district fixed effects and baseline outcomes significantly increases the precision of the estimate. At three months after the beginning of the intervention, the students in the treatment group were 5 to 7 percentage points more likely to abstain from tobacco than those in the control group, representing a 6 to 9 percent increase from baseline. These results suggest the 'top of mind' reminder from bi-weekly surveillance and commitment contract are effective in the short-term. We find that this effect persists at eight months after the beginning of the intervention, when regular monitoring and reporting no longer took place. At 8 months after the beginning of the intervention, students in the treatment group are 5 percentage points more likely to abstain from tobacco. The persistence of the effect suggests that the effect cannot be attributed to the weekly surveillance alone.

We then compare individual commitment to school competition (Panel B). We find that the individual commitment drives the program effect. At 3 months after the start of the intervention, the adjusted models show that the individual commitment is associated with a 5 to 7 percentage point higher probability of tobacco abstinence. The estimated effect is similar for the competition arm. These results show that school competition has no additional effect on smoking abstinence. This result may be due to several reasons. First, in spite of the peer pressure associated with tobacco use among adolescents, smoking is ultimately carried out by the individual student and group-based prevention may be less effective in this setting. Second, the 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 the norm in schools.

Table 2 presents the effects of the program using a difference-in-differences estimator. The treatment is associated with a 5 percentage point increase in the probability of abstaining from tobacco use (Panel A). The estimated effect remains largely unchanged when we include individual and class baseline outcomes and individual characteristics. The individual pledge appears to be just as effective in preventing tobacco use compared to the school competition (Panel B), which is consistent with our 3 and 8-month results. These results are consistent with the individual pledge as the main driver of the treatment effect. It is unclear whether the school competition had no additional effect on tobacco abstinence due to limited social effects (by emphasizing group identity) in this setting or the lack of individual incentives under school competition.

5.2 Peer effects

We explore the role of students' peer group since smoking is a social activity. We explore the role of friends and peers who smoke as sources of heterogeneity and exploit the random assignment in 8th grade as a source of variation to analyze peer effects.

We create an index for students' peers in 7th grade to explore heterogeneity by peers at baseline. Students were asked to list their 5 closest friends in their 7th grade class at baseline. On average, they have 1.5 friends who smoke, with a median of 2. On average, respondents believe their friends would smoke with a probability of 1.25 out of 10. We use the share of 7th grade students treated in the school, the share of the class with verified abstinence at baseline, the number of non-smokers among their 5 closest friends, and the average probability their 5 closest friends would not smoke. We create z-scores for each item, and take the average score for our 7th grade index. We create an indicator for high performance in 7th grade if the index is above the median. We interact the treatment indicator with the continuous index and the indicator for high performance. We find no heterogeneity by students' peer performance in 7th grade analyses (Table 3). The estimated effect of the treatment is similar to our earlier estimate (Panel A, col. 6), and the program effect is driven by the individual commitment contracts (Panel B, col. 6).

We also explore individual components of the index. We include an interaction term that takes the value one if at least two of the individual's five closest friends in the school smoked. Having more friends who smoke has no differential effect on the students' response to the intervention at three and eight months after the beginning of the intervention (Table B.4). The program effect is similar to our earlier estimate. We also explore a lower threshold of at least 1 friend who smoked at baseline and find similar result under difference-in-differences. Additionally, we ask students to estimate the probability that their 5 closest friends would smoke in the next 3 months. We interact the treatment indicator and an indicator that takes the value one if students believe none of their friends would begin smoking. We also find no differential response, but the main treatment effect remains positive (Table B.5). Similarly, we find no differential effect by friends' above median probability of smoking (this corresponds to 1.25 out of 10).

Our pilot was implemented in at most three 7th grade classes in the school. We explore the role of program saturation by interacting the treatment indicator with the share of 7th grade students who were treated. We include an indicator that takes the value one if the share in treatment is above the median, which corresponds to 75% of the school. We find no differential program effect on the students' response in high and low-saturation schools (Table

B.6).[23] The program effect is similar to our earlier estimate under difference-in-differences.

We leverage the random assignment of students in the new academic year when they enter 8th grade. We create an index for students' peers in 8th grade. It is possible that students are more likely to abstain from tobacco use when more of their classmates were treated in the previous year. We use the share of the 8th grade class who were treated (while they were in 7th grade) and the 8th grade class share of verified tobacco abstinence when the students were in 7th grade (based on the verified abstinence at the audit). We use z-scores to create the index and create an indicator for above median values. We find evidence of peer effects under difference-in-differences (Table 4).The results suggest that students perform better when surrounded by classmates who abstain from smoking. Interestingly, we find that peer effects appear stronger under school competition (Panel B).

We explore each component by including separate interaction terms. We include an interaction term that takes the value one if the share of 8th grade students who were treated in the previous year is above the median, which corresponds to 75% of the class. We find no differential effect by the share of the 8th class in treatment, and the main treatment effect is no longer statistically significant when students are in 8th grade (Table B.7). One plausible explanation is due to student performance at baseline and the 3-month follow-up. We include an interaction term that takes the value one if the share of verified abstinence is above the median at baseline and at the 3-month follow up, these correspond to 75%. We find a differential effect by the share of the 7th grade class who successfully abstained at baseline under difference-in-differences, but we find no differential effect by the share of the 8th grade class who abstained (Table B.9). These results suggest that the role of peer effects and peer pressure in smoking initiation may be limited in early adolescence, but changes in the circle of friends may play a role in a teen's decision to smoke.

5.3 Effects on related outcomes

We explore additional outcomes that relate to the intensity of smoking (Table 5). We explore self-reported frequency of smoking by combining students' cotinine test and self-reported last smoked in 2 weeks before the survey, and one month before the survey. We find that the program is effective in increasing self-reported tobacco abstinence in 2 weeks and one month. These results are consistent with our main outcome of interest of tobacco abstinence in the few days prior to the cotinine test. We also use the number of friends who report on a student as a measure of attitude towards smoking openly among their friends. We find that the program is associated with a 1.7 fewer students being reported as a smoker. This suggests

[23] Using an indicator for large schools with more than three 7th grade classes yields similar results.

that even if students continued to smoke, they were less likely to do so in close proximity to the school or their peers from school.

We also explore outcomes that relate to tobacco use: tobacco knowledge, smoking intentions, and students' belief that their parents would find smoking unacceptable in Table 6. Students on average correctly answered 5.8 tobacco-related questions (out of 9) at baseline. The treatment has no statistically significant effect on students' knowledge on the risks associated with tobacco use. The treatment also has no statistically significant effect on students' smoking intentions, although most students who failed to abstain had not intended to smoke. This result further suggests the role of time inconsistency and subjective belief among adolescents. We include truancy as a related disciplinary measure and find no significant treatment effect.

The treatment intended to increase parental involvement by inviting both parents and students to sign the agreement. Based on our pre-intervention conversations with parents, they believed that it was too early to begin discussing tobacco use with their 7th graders. In addition, 23% of students are baseline did not know whether their parents would find smoking acceptable. The treatment sought to change the norm partly through parental attitude. Unfortunately, the treatment has no statistically significant effect on students' belief of parental attitude towards tobacco use (Table B.8). We also explore whether students in the treatment group are more likely to discuss tobacco use with their parents and friends. We find a reduction in discussion with their friends and no significant effect on discussion with parents. It is therefore unclear whether the program increased parental engagement on the topic.[24]

5.4 Heterogeneous treatment effects

Student characteristics We explore heterogeneity by student characteristics. The index is based on z-scores of students' own intention to not smoke, never smoking, tobacco knowledge, time consistency, and risk aversion. At baseline, students were asked to choose a certain option and several gambles that correspond to different risk preference parameters. Students were also asked several questions about waiting for a hypothetical monetary gift that allows us to explore the role of students' time preference, including time consistency. The treatment has no effect on students' intention to smoke, but students' intention to smoke may be a source of heterogeneity, so we include this in the index. We find no significant heterogeneity by students' characteristics at baseline, and estimated effect of the treatment is similar to our earlier estimate (Table 7).

[24]We are conducting a follow-up study to target parents.

We also explore individual components of the index. Students who engage in risky behavior such as smoking may have a different appetite for risk compared to those who do not smoke, and such students may respond differently to the intervention. We interact the treatment indicator with students' choice of the certain outcome since almost 40% of students chose the certain option at baseline (Table B.10). We find no differential effect among students who chose certainty over a gamble, suggesting that risk preference among adolescents may not be as influential in students' decision to smoke in this setting.

Similarly, students who exhibit impatience or time inconsistency may respond differentially to the intervention. About 70% of students exhibit time inconsistency, and almost 15% exhibit impatience and patience, so we interact the treatment indicator with each time preference (Table B.11).[25] We find no differential effect among students who exhibited impatience. Interestingly, we find no differential effect among students who exhibited time inconsistency. The hypothetical nature of the question may have differed from their actual decision-making process.

To explore students' intention to smoke, we include an interaction term that takes the value one if students placed a positive probability on smoking by the end of the academic year (Table B.13). We find no differential effect among students who believed they might smoke in the short-term. This suggests that adolescents may have difficulty predicting their future decisions as their executive function is still developing.

We also explore whether the intervention is more beneficial to students who had not been smoking at baseline. We interact the treatment variable with an indicator that takes the value one if the student did not smoke in the past 2 weeks (past month) and passed the biochemical verification (Table B.12). We also interact the treatment variable with the indicator for ever smoked at baseline. We find no heterogeneity by the last time a student smoked, including heterogeneity by ever smoking, which suggests that the intervention may have benefited students who were occasional smokers at baseline.

Household characteristics We explore the role of household characteristics on children's response to the intervention. A particularly important channel is parental smoking status since parents transmit their attitude towards tobacco to their children. At three months after the start of the program, students whose parents smoke are less likely to abstain, but the interaction is no longer statistically significant at eight months after the beginning of the intervention (Table B.14). The estimated effect is similar under difference-in-differences at 8 months. These results highlight the importance of parents as role models, at least in the

[25]Interacting the treatment with time inconsistency and impatience to explore the differential effect relative to those who show patience yields similar results.

short-term.

For all other characteristics, we create an index based on the following z-scores: non-low maternal education, non-low paternal education, household size, land ownership, and a count of non-land asset ownership. We create an indicator that takes the value one for above median values.[26] We find no significant heterogeneity by household characteristics (Table 8), which suggests that the intervention does not appear to benefit students with particularly low or high socio-economic backgrounds.

School characteristics School quality may be a source of heterogeneity in this context. Specifically, schools that are more academically demanding may implement the program more strictly, or the students are more academically oriented and are less interested in smoking. A related influence may be the number of teachers who smoke in the school. Another source of heterogeneity may be the location of the school since smoking is more prevalent in rural areas in Indonesia. We create an index based on the z-scores of the following: no teachers smoking, Indonesian language test score, mathematics test score, urban location, distance from the nearest treatment, the number of 7th grade classes in the school. We then create an indicator that takes the value one if the index is above the median value. We find no significant heterogeneity by school characteristics (Table 9).

We also explore separate components of the index. We interact the treatment indicator with an indicator that takes the value one if the school's average mathematics score on the national examination is above the median (Table B.17). We find that schools with higher test scores are more likely to benefit from the intervention at 3 and 8 months after the beginning of the intervention. About 30% of schools report no smoking among teachers (Table B.16). We find no heterogeneity in treatment effects by teachers' smoking status, and the main treatment effect remains positive at 3 and 8 months intervention.[27] We find no differential effect by urban location, suggesting that smoking among adolescents may be similar across urban and rural areas (Table B.18). We also find no differential effect by distance to the nearest treated school, suggesting that treatment spillover did not drive our earlier results.

[26] We also explore individual household characteristics: maternal and paternal education and the occupational status of the head of the household (Table B.15). Even though maternal education is often linked to better health outcomes for children, we find limited evidence in this case. We also find no differential effect by paternal education. Under difference-in-differences, children whose parents have low occupational status appear to benefit more from the program.

[27] Anecdotally, student council members suggested that teachers' who do not smoke are more supportive of the implementation of the program.

5.5 Alternative outcomes

We explore alternative definitions of tobacco use for robustness. First, we restrict the sample to students with non-missing cotinine test and only use the test as an outcome variable (Table B.19, Panel A). Second, we supplement the biochemically verified test with friends' reports (Panel B). Third, we assume students who did not take the cotinine test as non-smokers if they were non-smokers at baseline (Panel C). The estimated program effects are similar under these alternative outcome measures, suggesting the effectiveness of the treatment in preventing smoking initiation.

5.6 Cost comparison

We compare the cost of our prevention program 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 uses a monetary commitment contract is associated with a 14% increase in cessation at 6 months, and the program effect is 5% at 12 months, about a third of the effect size at 6 months [Giné et al., 2010]. The program cost per quit is $700.

The costs associated with our prevention program include the unannounced cotinine tests (twice per year, 50% of students selected in 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 cotinine test kit costs $2 per test, printing costs would be $0.10. Administrative costs would be $6000, assuming a $500 monthly administrative costs. The cost of the program is approximately $4.40 per student. Our prevention program effect at 6 months is 6%, and we assume that the program effect would be a third in 12 months, at 2%. Based on the expected 12-month effect, the cost per smoker averted would be $220, which is much lower than the cost for cessation. The program effect can be as low as $0.6% (at 12 months) for the prevention cost to be comparable to the cost of smoking cessation.

6 Conclusion

We find that an intervention that involves non-monetary incentives lowers the probability of smoking among school children in a high prevalence setting. We find that parental influence matters in the short-term, but not in the longer-term. More importantly, we find no heterogeneity by student characteristics, suggesting that the intervention does not favor students with particular risk preference or socio-economic background. The intervention is particularly effective in high quality schools, suggesting the important role of the school

environment.

Broadly speaking, the intervention suggests that commitment contracts work even among adolescents, for whom self-control tends to be a concern. Moreover, this intervention suggests that tobacco control programs can be low-cost and easily scalable, even in a high prevalence setting.

References

Dan Acland and Vinci Chow. Self-control and demand for commitment in online game playing: evidence from a field experiment. *Journal of the Economic Science Association*, 4(1):46–62, Jul 2018. ISSN 2199-6784. doi: 10.1007/s40881-018-0048-3. URL `https://doi.org/10.1007/s40881-018-0048-3`.

Dan Ariely and Klaus Wertenbroch. Procrastination, deadlines, and performance: Self-Control by precommitment. *Psychological Science*, 13(3):219–224, May 2002. ISSN 09567976. URL `http://www.jstor.org/stable/40063710`. ArticleType: primary_article / Full publication date: May, 2002 / Copyright © 2002 Association for Psychological Science.

Liang Bai, Benjamin Handel, Edward Miguel, and Gautam Rao. Self-control and demand for preventive health: Evidence from hypertension in india. Working Paper 23727, National Bureau of Economic Research, August 2017. URL `http://www.nber.org/papers/w23727`.

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

Sarah Barber, 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.

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

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

David Card and Laura Giuliano. Peer effects and multiple equilibria in the risky behavior of friends. *Review of Economics and Statistics*, 95(4):1130–1149, 2013/12/10 2013. doi: 10.1162/REST{_}a{_}00340. URL `http://dx.doi.org/10.1162/REST_a_00340`.

Joseph R DiFranza, 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*, 120(4):e974–e983, 2007.

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

Susan T Ennett, 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*, 63(1): 48–62, 2001.

Emma L Giles, 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*, 9(3):e90347, 2014.

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

Saji S. Gopalan, 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*, 100(0):72–83, 1 2014. doi: http://dx.doi.org/10.1016/j.socscimed.2013.10.030. URL `http://www.sciencedirect.com/science/article/pii/S0277953613005789`.

Scott D. Halpern, 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*, 372(22):2108–2117, 2015. doi: 10.1056/NEJMoa1414293. URL `http://dx.doi.org/10.1056/NEJMoa1414293`. PMID: 25970009.

Todd F. Heatherton, 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*, 86(9):1119–1127, 1991. doi: 10.1111/j.1360-0443.1991.tb01879.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1360-0443.1991.tb01879.x`.

Kosuke Imai, 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*, 24(1):29–53, 2009.

Ahmed Jamal, 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*, 66(23):597, 2017.

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

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

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

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

Eleni Mantzari, 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*, 75:75 – 85, 2015. ISSN 0091-7435. doi: http://dx.doi.org/10.1016/j.ypmed.2015.03.001. URL `http://www.sciencedirect.com/science/article/pii/S0091743515000729`.

Nawi Ng, 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*, 16(3):190–196, June 2007a. doi: 10.1136/tc.2006.018895. URL `http://tobaccocontrol.bmj.com/cgi/content/abstract/16/3/190`.

Nawi Ng, 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*, 22(6):794–804, 2007b.

Mark Nichter. Introducing tobacco cessation in developing countries: an overview of project quit tobacco international. *Tobacco Control*, 15(Suppl_1):i12–17, June 2006. doi: 10.1136/tc.2005.014704. URL `http://tobaccocontrol.bmj.com/cgi/content/abstract/15/suppl_1/i12`.

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

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

Matthew Rabin. Healthy habits: Some thoughts on the role of public policy in healthful eating and exercise under limited rationality. In Adam Oliver, editor, *Behavioural Public*

Policy, chapter 5, pages 115–139. Cambridge University Press, 2013.

Marissa B Reitsma, 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*, 389(10082):1885–1906, 2017.

Heather Royer, 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*, 7(3):51–84, 2015.

Frank Schilbach. Alcohol and self-control: A field experiment in india. *American Economic Review*, Forthcoming.

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

Steve Sussman, James Miyano, Louise Ann Rohrbach, Clyde W. Dent, and Ping Sun. Six-month and one-year effects of project ex-4: A classroom-based smoking prevention and cessation intervention program. *Addictive Behaviors*, 32(12):3005–3014, 12 2007. doi: http://dx.doi.org/10.1016/j.addbeh.2007.06.016. URL `http://www.sciencedirect.com/science/article/pii/S0306460307001748`.

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

Roger E. Thomas, Julie McLellan, and Rafael Perera. School-based programmes for preventing smoking. *Evidence-Based Child Health: A Cochrane Review Journal*, 8(5): 1616–2040, 2013. ISSN 1557-6272. doi: 10.1002/ebch.1937. URL `http://dx.doi.org/10.1002/ebch.1937`.

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

Robert West, Peter Hajek, Lindsay Stead, and John Stapleton. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction*, 100(3):299–303, 2005. ISSN 1360-0443. doi: 10.1111/j.1360-0443.2004.00995.x. URL `http://dx.doi.org/10.1111/j.1360-0443.2004.00995.x`.

World Health Organization. Who global report: Mortality attributable to tobacco. Technical report, World Health Organization, Geneva, Switzerland, 2012a.

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

World Health Organization. Global adult tobacco survey (GATS): Indonesia report 2011. *Jakarta: WHO Regional Office for South East Asia*, 2012c.

World Health Organization. WHO global report on trends in prevalence of tobacco smoking. Technical report, World Health Organization, Geneva, Switzerland, 2015a.

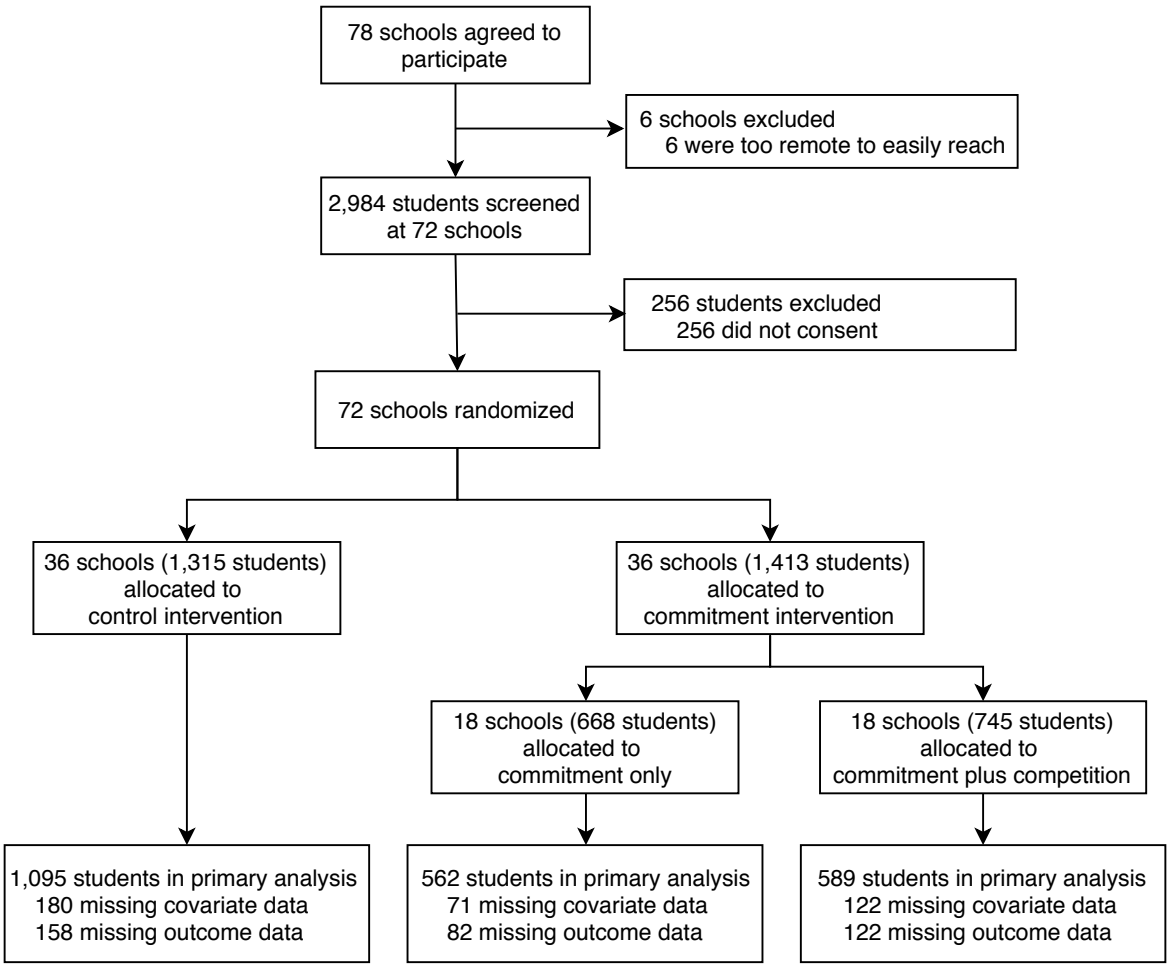
World Health Organization. Global youth tobacco survey: Indonesia report 2014. Technical report, World Health Organization Regional Office for Southeast Asia, New Delhi, India, 2015b.

World Health Organization. *WHO report on the global tobacco epidemic 2015: raising taxes on tobacco*. World Health Organization, 2015c.

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

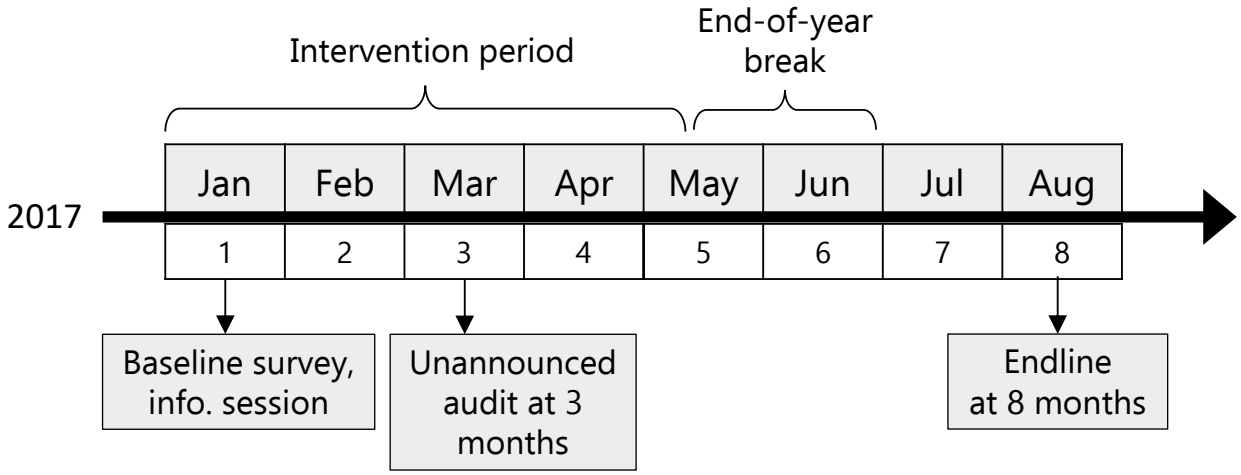
Tables and Figures

Figure 1: Study flowchart



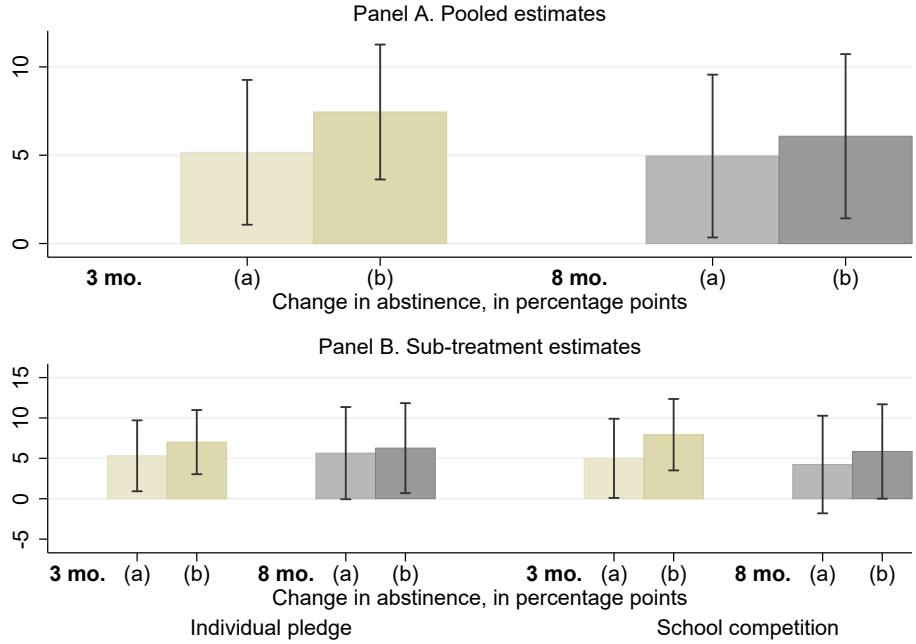
Notes: Observations with missing covariate data and missing outcome data are not mutually exclusive.

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.

Figure 3: Intervention effects on abstinence at 3 and 8 months



Notes: This figure shows the intent-to-treat effect of the pooled intervention (Panel A) and sub-treatment interventions (Panel B) on smoking abstinence at 3 and 8 months, expressed in percentage points. Model (a) includes district fixed effects, each student's baseline abstinence, and class average abstinence at baseline (excluding the student). Model (b) adds the school's average math test score, number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, and parental smoking status. Error bars are 95% confidence intervals. Robust standard errors are clustered by school. Significance: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Table 1: Baseline characteristics and balance under pair matching

	Control	Treatment	t-test	Adjusted difference	Individual	Competition
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. School						
Distance	17.431	17.601	0.93	0.657	0.926	0.448
	(10.647)	(9.560)		(1.512)	(1.638)	(1.976)
Number of students	374.179	362.021	0.73	-8.920	18.235	-35.300
	(149.382)	(147.852)		(39.809)	(53.467)	(43.255)
Student-to-teacher ratio	0.073	0.076	0.78	0.009	0.019	0.000
	(0.029)	(0.044)		(0.015)	(0.028)	(0.011)
Number of student council	11.343	9.164	0.61	-1.687	-1.010	-3.002
members	(10.859)	(10.743)		(2.331)	(3.570)	(1.929)
Average math score	57.122	52.525	0.26	-3.962	-3.843	-3.674
	(14.483)	(13.292)		(3.467)	(4.576)	(3.986)
N	36	36				
Joint significance test (p-value)				0.316	0.364	0.119
Panel B. Individual						
Abstained in the past 7 days	0.788	0.775		-0.010	-0.027	0.007
	(0.409)	(0.418)		(0.034)	(0.043)	(0.037)
Class average: Abstained in	0.784	0.771		-0.010	-0.028	0.013
the past 7 days	(0.187)	(0.168)		(0.034)	(0.044)	(0.036)
Household size	4.748	4.705		-0.081	-0.318**	0.127
	(1.651)	(1.558)		(0.153)	(0.142)	(0.211)
Land ownership	0.551	0.509		-0.052	-0.071	-0.039
	(0.498)	(0.500)		(0.032)	(0.043)	(0.040)
Head of household in	0.387	0.339		-0.063*	-0.070	-0.056
skilled occupation	(0.487)	(0.474)		(0.037)	(0.057)	(0.036)
Parent smokes	0.443	0.476		0.043	0.039	0.052
	(0.497)	(0.500)		(0.028)	(0.030)	(0.036)
Low maternal education	0.135	0.131		0.004	-0.001	0.006
	(0.341)	(0.337)		(0.021)	(0.027)	(0.026)
Low paternal education	0.151	0.137		-0.010	0.017	-0.036
	(0.359)	(0.344)		(0.026)	(0.038)	(0.027)
Joint significance test (p-value)				0.304	0.229	0.182

Notes: Standard deviations are in parentheses in Columns 1 and 2. Column 3 reports the p-values from t-tests of the difference in means between the treatment and control groups. Each cell in Columns 4-6 is the coefficient on the treatment variable from a separate OLS regression with district fixed effects. Standard errors are clustered at the school level.

Table 2: Program effect on tobacco abstinence: Difference-in-differences at 8 months

	(1)	(2)	(3)
Panel A. Pooled estimates			
Treatment	-0.011	0.001	0.011
	(0.035)	(0.009)	(0.011)
Treatment × Post	0.044**	0.047**	0.047**
	(0.022)	(0.022)	(0.022)
Obs.	6059	6059	6059
Dep. variable mean for control group	0.773		
Pseudo R-sq.	0.00	0.38	0.39
Class baseline outcome	No	Yes	Yes
Additional controls	No	No	Yes
Panel B. Sub-treatment estimates			
Commitment × Post	0.042*	0.045*	0.044*
	(0.024)	(0.024)	(0.024)
School competition x Post	0.046	0.049*	0.049*
	(0.030)	(0.029)	(0.029)
Commitment	-0.027	0.006	0.012
	(0.042)	(0.011)	(0.013)
Competition	0.006	-0.004	0.010
	(0.039)	(0.011)	(0.013)
Obs.	6059	6059	6059
Dep. variable mean for control group	0.773		
Pseudo R-sq.	0.00	0.38	0.39
Class baseline outcome	No	Yes	Yes
Additional controls	No	No	Yes

Notes: Post takes the value one for the follow-up period. Treatment (Commitment, Competition) takes the value one if the school is randomized into treatment (the individual commitment arm, school competition arm). Col. 1 includes district fixed effects. Col. 2 includes district fixed effects and baseline outcome. Col. 3 adds: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table 3: Program effect on tobacco abstinence by 7th grade performance

	(1)	(2)	(3)	(4)	(5)	(6)
	Index			High index		
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Panel A. Pooled estimates						
Treatment	0.060	0.072*	0.026	0.059**	0.047*	-0.010
	(0.048)	(0.043)	(0.035)	(0.024)	(0.027)	(0.016)
High performance	0.046***	0.044***	0.086***	0.052**	0.031	0.118***
	(0.013)	(0.012)	(0.009)	(0.025)	(0.025)	(0.017)
Treatment ×	0.009	-0.003	-0.011	0.056	0.041	0.014
High performance	(0.017)	(0.015)	(0.013)	(0.034)	(0.038)	(0.025)
Treatment × Post			0.093**			0.067***
			(0.038)			(0.026)
Treatment × Post ×			-0.010			0.005
High performance			(0.013)			(0.025)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.776
Pseudo R-sq.	0.26	0.23	0.16	0.25	0.23	0.15
Panel B. Sub-treatment estimates						
Commitment	0.047	0.069	-0.016	0.056**	0.050	-0.008
	(0.052)	(0.053)	(0.043)	(0.026)	(0.031)	(0.017)
High performance	0.047***	0.044***	0.086***	0.052**	0.030	0.118***
	(0.013)	(0.012)	(0.009)	(0.025)	(0.025)	(0.017)
Commitment ×	0.011	-0.001	0.009	0.047	0.038	0.038
High performance	(0.018)	(0.020)	(0.016)	(0.045)	(0.055)	(0.034)
Competition	0.074	0.076	0.059	0.062**	0.043	-0.013
	(0.058)	(0.049)	(0.040)	(0.028)	(0.036)	(0.022)
Competition×	0.006	-0.005	-0.026*	0.065*	0.045	-0.004
High performance	(0.021)	(0.016)	(0.014)	(0.037)	(0.039)	(0.030)
Commitment × Post			0.127**			0.068***
			(0.054)			(0.025)
Competition× Post			0.066			0.067*
			(0.048)			(0.036)
Commitment × Post ×			-0.027			-0.021
High performance			(0.020)			(0.042)
Competition × Post ×			0.003			0.026
High performance			(0.016)			(0.029)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.761	0.815	0.790	0.761
Pseudo R-sq.	0.26	0.23	0.16	0.25	0.23	0.15

Notes: The index is based on z-scores of the share of the school's 7th grade male students who were treated and the share of verified tobacco abstinence in each class. High performance takes the value one if the index is above the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table 4: Program effect on tobacco abstinence by 8th grade performance

	(1)	(2)	(3)	(4)
	Index		High index	
	8-month	Diff-in-diff	8-month	Diff-in-diff
Panel A. Pooled estimates				
Treatment	0.056**	-0.022	0.051*	0.014
	(0.024)	(0.014)	(0.027)	(0.019)
Performance	0.022	0.043***	0.049	0.055***
	(0.016)	(0.012)	(0.032)	(0.018)
Treatment ×	0.019	-0.056***	0.013	-0.066***
Performance	(0.021)	(0.015)	(0.042)	(0.025)
Treatment × Post		0.082***		0.050*
		(0.024)		(0.027)
Treatment × Post ×		0.056***		0.073**
Performance		(0.018)		(0.033)
Obs.	2011	8579	2011	8579
Dep. variable mean for control group	0.800	0.785	0.800	0.785
Pseudo R-sq.	0.22	0.13	0.22	0.13
Panel B. Sub-treatment estimates				
Commitment	0.058*	-0.009	0.046	0.024
	(0.031)	(0.015)	(0.030)	(0.021)
Performance	0.022	0.042***	0.048	0.054***
	(0.016)	(0.012)	(0.032)	(0.018)
Commitment ×	0.016	-0.052***	0.027	-0.058**
Performance	(0.029)	(0.019)	(0.043)	(0.028)
Competition	0.054*	-0.035*	0.056*	0.005
	(0.029)	(0.018)	(0.033)	(0.024)
Competition×	0.020	-0.060***	-0.003	-0.076**
Performance	(0.024)	(0.018)	(0.059)	(0.032)
Commitment × Post		0.067***		0.031
		(0.024)		(0.028)
Competition× Post		0.098***		0.067*
		(0.032)		(0.039)
Commitment × Post ×		0.046*		0.078**
Performance		(0.027)		(0.037)
Competition × Post ×		0.067**		0.072
Performance		(0.029)		(0.057)
Obs.	2011	8579	2011	8579
Dep. variable mean for control group	0.800	0.766	0.800	0.766
Pseudo R-sq.	0.22	0.13	0.22	0.13

Notes: The index is based on z-scores of the share of the 8th grade class who were treated and the 8th grade class share of verified tobacco abstinence when students were in 7th grade (in March). High performance takes the value one if the index is above the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table 5: Treatment effects on smoking intensity

	(1) Last smoked > 2 weeks ago	(2) Last smoked > 1 month ago	(3) Number of friends who reported
Treatment	0.009	-0.002	-1.724***
	(0.009)	(0.013)	(0.579)
Treatment × Post	0.053**	0.051**	
	(0.024)	(0.024)	
Obs.	4580	4580	2281
Dep. var. mean for control group	0.769	0.715	5.380
Pseudo R-sq.	0.46	0.37	0.15

Notes: District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table 6: Treatment effects on related outcomes

	(1) Knowledge	(2) Smoking intention	(3) Truancy
Treatment	0.152	0.055	0.007
	(0.132)	(0.093)	(0.007)
Treatment × Post	0.022	-0.115	-0.004
	(0.232)	(0.131)	(0.022)
Obs.	4691	4569	4569
Dep. variable mean for control group	5.784	1.132	0.037
Pseudo R-sq.	0.02	0.16	0.11

Notes: District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table 7: Program effect on tobacco abstinence by individual characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
		Index			High index	
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Panel A. Pooled estimates						
Treatment	0.075***	0.062**	-0.008	0.075***	0.067**	-0.008
	(0.021)	(0.025)	(0.015)	(0.024)	(0.029)	(0.016)
Characteristic index	0.109***	0.162***	0.208***	0.087***	0.134***	0.184***
	(0.026)	(0.028)	(0.018)	(0.023)	(0.020)	(0.015)
Treatment ×	0.019	-0.008	0.006	-0.001	-0.020	-0.011
Characteristic index	(0.038)	(0.039)	(0.027)	(0.034)	(0.032)	(0.023)
Treatment × Post			0.067***			0.067**
			(0.025)			(0.027)
Treatment × Post ×			0.004			-0.000
Characteristic index			(0.030)			(0.028)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.776
Pseudo R-sq.	0.26	0.26	0.19	0.25	0.25	0.18
Panel B. Sub-treatment estimates						
Commitment	0.068**	0.062	-0.019	0.068***	0.058*	-0.007
	(0.029)	(0.038)	(0.016)	(0.023)	(0.031)	(0.016)
Characteristic index	0.087***	0.134***	0.184***	0.109***	0.162***	0.208***
	(0.023)	(0.020)	(0.015)	(0.026)	(0.028)	(0.018)
Commitment ×	-0.001	-0.009	0.036	0.024	0.010	0.050
Characteristic index	(0.044)	(0.044)	(0.027)	(0.050)	(0.052)	(0.032)
Competition	0.082***	0.072**	0.002	0.082***	0.065**	-0.011
	(0.027)	(0.034)	(0.021)	(0.025)	(0.032)	(0.019)
Competition×	0.001	-0.031	-0.056*	0.014	-0.023	-0.032
Characteristic index	(0.037)	(0.034)	(0.029)	(0.039)	(0.042)	(0.032)
Commitment × Post			0.076**			0.065**
			(0.030)			(0.025)
Competition× Post			0.060*			0.070**
			(0.036)			(0.033)
Commitment × Post ×			-0.038			-0.034
Characteristic index			(0.040)			(0.043)
Competition × Post ×			0.035			0.037
Characteristic index			(0.035)			(0.041)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.761	0.815	0.790	0.761
Pseudo R-sq.	0.25	0.25	0.18	0.26	0.26	0.20

Notes: The index is based on z-scores of students' own intention to not smoke, never smoking, tobacco knowledge, time consistency, and risk aversion. High index takes the value one if the index is above the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table 8: Program effect on tobacco abstinence by household characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Index			High index		
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Panel A. Pooled estimates						
Treatment	0.074***	0.061**	-0.010	0.077***	0.071***	-0.007
	(0.019)	(0.024)	(0.014)	(0.026)	(0.027)	(0.019)
Characteristic index	-0.066*	0.017	-0.021	-0.051	-0.017	-0.011
	(0.037)	(0.038)	(0.028)	(0.033)	(0.025)	(0.024)
Treatment ×	0.025	-0.006	0.003	-0.003	-0.019	-0.005
Characteristic index	(0.029)	(0.031)	(0.025)	(0.034)	(0.028)	(0.027)
Treatment × Post			0.068***			0.082***
			(0.024)			(0.027)
Treatment × Post ×			-0.011			-0.029
Characteristic index			(0.021)			(0.022)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.776
Pseudo R-sq.	0.24	0.22	0.13	0.24	0.22	0.13
Panel B. Sub-treatment estimates						
Commitment	0.067***	0.062**	0.001	0.084***	0.084**	0.005
	(0.020)	(0.030)	(0.015)	(0.028)	(0.036)	(0.021)
Characteristic index	-0.069*	0.015	-0.022	-0.051	-0.018	-0.012
	(0.038)	(0.038)	(0.028)	(0.033)	(0.025)	(0.024)
Commitment ×	-0.003	-0.023	0.005	-0.027	-0.042	-0.005
Characteristic index	(0.034)	(0.037)	(0.030)	(0.037)	(0.032)	(0.032)
Competition	0.080***	0.058*	-0.020	0.070**	0.057*	-0.020
	(0.023)	(0.030)	(0.018)	(0.031)	(0.031)	(0.024)
Competition×	0.051	0.011	0.005	0.022	0.004	-0.002
Characteristic index	(0.032)	(0.040)	(0.027)	(0.038)	(0.036)	(0.030)
Commitment × Post			0.057**			0.085***
			(0.024)			(0.028)
Competition× Post			0.075**			0.079**
			(0.033)			(0.039)
Commitment × Post ×			-0.045*			-0.054**
Characteristic index			(0.026)			(0.025)
Competition × Post ×			0.015			-0.007
Characteristic index			(0.031)			(0.035)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.761	0.815	0.790	0.761
Pseudo R-sq.	0.24	0.22	0.13	0.24	0.22	0.13

Notes: The index is based on z-scores of non-low maternal education, non-low paternal education, household size, land ownership, and a count of non-land asset ownership. High index takes the value one if the index is above the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table 9: Program effect on tobacco abstinence by school characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Index			High index		
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Panel A. Pooled estimates						
Treatment	0.081***	0.070***	0.002	0.080***	0.047	-0.012
	(0.019)	(0.025)	(0.014)	(0.026)	(0.033)	(0.021)
Characteristic index	0.058	0.034	0.042*	0.032	0.038	0.023
	(0.037)	(0.041)	(0.023)	(0.033)	(0.036)	(0.021)
Treatment ×	-0.025	0.031	0.025	-0.012	0.036	0.006
Characteristic index	(0.033)	(0.044)	(0.026)	(0.039)	(0.053)	(0.028)
Treatment × Post			0.059**			0.071***
			(0.025)			(0.025)
Treatment × Post ×			-0.042			-0.010
Characteristic index			(0.028)			(0.037)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.776
Pseudo R-sq.	0.24	0.22	0.13	0.24	0.22	0.13
Panel B. Sub-treatment estimates						
Commitment	0.079***	0.083***	0.011	0.049*	0.010	-0.012
	(0.020)	(0.027)	(0.014)	(0.028)	(0.039)	(0.021)
Characteristic index	0.060	0.043	0.046**	0.032	0.037	0.024
	(0.037)	(0.039)	(0.023)	(0.032)	(0.035)	(0.021)
Commitment ×	0.000	0.098*	0.022	0.035	0.102*	0.023
Characteristic index	(0.036)	(0.052)	(0.027)	(0.040)	(0.054)	(0.029)
Competition	0.079***	0.055	-0.007	0.100***	0.073*	-0.011
	(0.024)	(0.035)	(0.018)	(0.027)	(0.037)	(0.025)
Competition×	-0.055	-0.052	0.021	-0.061	-0.031	-0.022
Characteristic index	(0.039)	(0.060)	(0.038)	(0.053)	(0.082)	(0.038)
Commitment × Post			0.062***			0.049*
			(0.023)			(0.027)
Competition× Post			0.055			0.087***
			(0.035)			(0.033)
Commitment × Post ×			-0.002			0.027
Characteristic index			(0.028)			(0.029)
Competition × Post ×			-0.086			-0.043
Characteristic index			(0.057)			(0.066)
Obs.	2161	2246	9376	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.761
Pseudo R-sq.	0.24	0.22	0.13	0.24	0.23	0.13

Notes: The index is based on z-scores of no teachers smoking, Indonesian language test score, mathematics test score, urban location, distance from the nearest treatment, the number of 7th grade classes in the school. High index takes the value one if the index is above the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

A Sample selection

A.1 Sample size calculation

We conducted sample size calculations for a cluster randomized trial with an individual-level binary outcome.[28] 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 a 5-percentage point magnitude 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 Jogjakarta. 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 sub-districts in Sleman.[29]

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.

[28] The 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.

[29] Some tobacco is also grown in the following sub-districts: Prambanan, Ngemplak, Sleman, Seyegan, and Tempel. Source: http://jogja.tribunnews.com/2015/04/19/dua-kecamatan-jadi-sentra-produksi-tembakau.

A.2 Treatment assignment

We used pair matching 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).

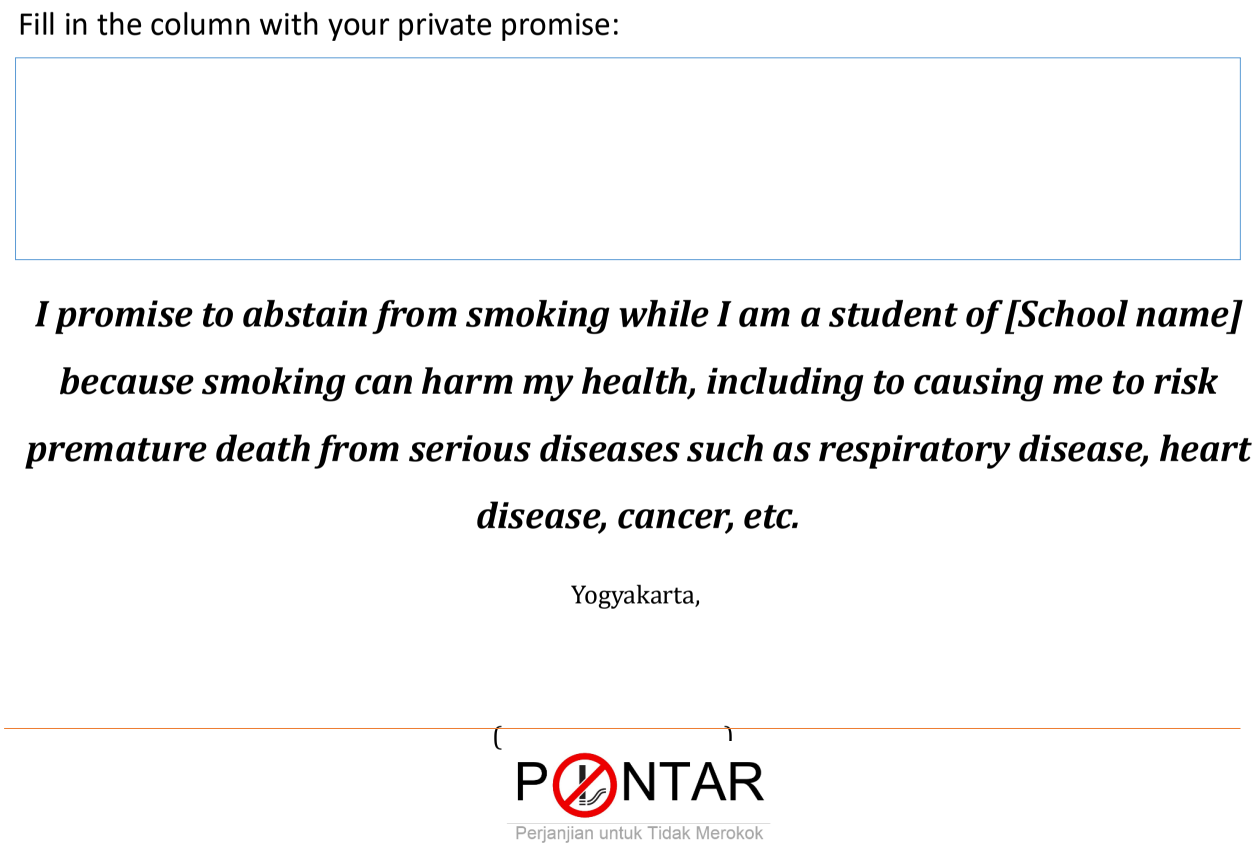
Our matching procedure included the following covariates: district, subdistrict, distance from the school to Yogyakarta proper, number of male students and classrooms in 7th grade, total number of students, teachers, classes, and student council members in the school, electricity capacity, and average 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.1: List of subdistricts

	Number of schools		
District	Control	Commitment	Commitment plus competition
Sleman	17	12	7
Kulon Progo	19	6	11
N	36	18	18

Figure B.1: Students' Commitment Contracts

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



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.2: Parents' Commitment Contract

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 tobacco use.

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

Table B.2: Balance in 8th grade

	(1)	(2)	(3)
			Adjusted difference
	Treatment	Control	P(same classsXtreated)
Abstained in the past 7 days	0.773	0.778	-0.046
(baseline)	(0.419)	(0.415)	(0.092)
Household size	4.74	4.70	-0.142***
	(1.66)	(1.58)	(0.050)
Land ownership	0.55	0.51	-0.033
	(0.50)	(0.50)	(0.104)
Head of household in	0.35	0.30	-0.093
skilled occupation	(0.48)	(0.46)	(0.056)
Parent smokes	0.44	0.47	0.086*
	(0.50)	(0.50)	(0.051)
Low maternal education	0.375	0.357	-0.006
	(0.484)	(0.479)	(0.091)
Low paternal education	0.409	0.396	0.009
	(0.492)	(0.489)	(0.113)
Joint significance test (p-value)			0.196

Notes: Each cell is the coefficient on the treatment variable from a separate OLS regression with district fixed effects. The dependent variable is an indicator that takes the value one if the student remains in the same class in 8th grade (ie. moving from 7A to 8A) and the school was assigned to the treatment group. Standard errors are clustered at the school level.

Table B.3: Program effect on tobacco abstinence

	(1)	(2)	(3)	(4)	(5)	(6)
	3-month			8-month		
	Unadjusted	Partially adjusted	Fully adjusted	Unadjusted	Partially adjusted	Fully adjusted
Panel A. Pooled estimates	0.038	0.052**	0.074***	0.037	0.049**	0.061**
Treatment	(0.037)	(0.021)	(0.019)	(0.038)	(0.023)	(0.024)
Obs.	2161	2161	2161	2246	2246	2246
Dep. var. mean for control group	0.815	0.815	0.815	0.790	0.790	0.790
R-sq.	0.00	0.21	0.24	0.00	0.20	0.22
Panel B. Sub-treatment estimates						
Commitment	0.014	0.053**	0.069***	0.020	0.057*	0.063**
	(0.045)	(0.023)	(0.020)	(0.047)	(0.031)	(0.030)
School competition	0.065	0.050*	0.080***	0.051	0.040	0.054*
	(0.043)	(0.025)	(0.023)	(0.051)	(0.030)	(0.031)
Obs.	2161	2161	2161			
Dep. var. mean for control group	0.815	0.815	0.815	2016	2016	2016
Pseudo R-sq.	0.00	0.21	0.24			

Notes: Partially adjusted includes district fixed effects, baseline smoking status. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.4: Program effect on tobacco abstinence by friends' smoking status

Variable	(1)	(2)	(3)	(4)	(5)	(6)
	At least 1 friend smoked			At least 2 friends smoked		
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Treatment	0.051*	0.031	-0.012	0.073***	0.049	-0.019
	(0.030)	(0.039)	(0.021)	(0.025)	(0.031)	(0.018)
Variable	-0.102***	-0.092***	-0.075***	-0.071***	-0.082***	-0.149***
	(0.028)	(0.033)	(0.016)	(0.024)	(0.022)	(0.017)
Treatment ×	0.037	0.045	0.009	0.010	0.030	0.035
Variable	(0.035)	(0.042)	(0.023)	(0.033)	(0.034)	(0.023)
Treatment × Post			0.049*			0.061**
			(0.029)			(0.027)
Treatment × Post× Variable			0.033			0.011
			(0.025)			(0.024)
Obs.	2161	2246	8654	2161	2246	9362
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.776
R-sq.	0.25	0.23	0.33	0.25	0.23	0.16

Notes: Each source of heterogeneity takes the value one if a student has at least (one) two friends who smoke at baseline. Two corresponds to the median value. District fixed effects included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.5: Program effect on tobacco abstinence by friends' predicted smoking status

	(1)	(2)	(3)	(4)	(5)	(6)
	None expected to smoke			High probability of smoking		
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Treatment	0.073***	0.063***	-0.006	0.086***	0.064**	0.029
	(0.019)	(0.024)	(0.015)	(0.029)	(0.033)	(0.025)
Variable	0.062***	0.040	0.130***	-0.046**	-0.034	-0.111***
	(0.022)	(0.029)	(0.017)	(0.023)	(0.023)	(0.015)
Treatment ×	0.007	-0.014	-0.024	-0.019	-0.006	-0.033
Variable	(0.031)	(0.041)	(0.026)	(0.033)	(0.034)	(0.029)
Treatment × Post			0.069***			-0.004
			(0.025)			(0.027)
Treatment × Post ×			-0.009			0.039
Variable			(0.026)			(0.026)
Obs.	2161	2246	9378	2155	2245	9347
Dep. variable mean for control group	0.815	0.790	0.776	0.816	0.790	0.777
Pseudo R-sq.	0.24	0.22	0.14	0.24	0.22	0.15

Notes: 'None expected to smoke' corresponds to 25th percentile of respondents' average probability of smoking among their friends. 'High probability of smoking' corresponds to higher than 0.125 average probability of smoking among their friends. District fixed effects included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.6: Program effect on tobacco abstinence by share of 7th grade treated

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Treatment	0.092***	0.097***	-0.007
	(0.032)	(0.034)	(0.018)
High share 7th grade	-0.002	0.038	-0.020
	(0.037)	(0.037)	(0.021)
Treatment ×	-0.034	-0.063	-0.010
High share 7th grade	(0.039)	(0.050)	(0.026)
Treatment × Post			0.070**
			(0.028)
Treatment × Post ×			-0.003
High share 7th grade			(0.032)
Obs.	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776
Pseudo R-sq.	0.24	0.22	0.13

Notes: High share treated corresponds to above median share of 7th grade treated in the school, corresponding to 75% of the 7the grade treated. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.7: Program effect on tobacco abstinence by share of 8th grade treated

	(1)	(2)
	8-month	Diff-in-diff
Treatment	-0.005	-0.010
	(0.026)	(0.015)
High share treated	-0.011	-0.000
	(0.033)	(0.020)
Treatment x High share treated	0.101**	0.038
	(0.042)	(0.026)
Treatment × Post		0.031
		(0.025)
Treatment × Post ×		0.004
High share treated		(0.028)
Obs.	2246	8654
Dep. variable mean for control group	0.826	0.818
Pseudo R-sq.	0.25	0.33

Notes: High share treated corresponds to above median share of 8th grade classmates treated, corresponding to 75% of the class treated. District fixed effects included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.8: Treatment effects on related outcomes

	(1) High Knowledge	Parents do not accept	(4) Discuss with parents	(5) Discuss with friends
Treatment	0.022	0.013	-0.016	-0.037*
	(0.031)	(0.025)	(0.014)	(0.022)
Treatment × Post	0.007	0.013		
	(0.042)	(0.024)		
Obs.	4691	3935	2224	2224
Dep. variable mean for control group	0.449	0.721	0.073	0.140
Pseudo R-sq.	0.01	0.06	0.04	0.04

Notes: High knowledge takes the value one if the student scored more than the median of 6 (out of 9) on tobacco knowledge. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table B.9: Program effect on tobacco abstinence by class performance

	(1)	(2)	(3)	(4)	(5)
		7th grade		8th grade	
	3-month	8-month	Diff-in-diff	8-month	Diff-in-diff
Panel A. 7th grade					
Treatment	0.074***	0.061**	-0.005	0.047	-0.011
	(0.019)	(0.024)	(0.014)	(0.033)	(0.016)
High performance	0.049	-0.002	0.068	0.025	0.030*
	(0.162)	(0.128)	(0.083)	(0.033)	(0.018)
Treatment ×	-0.010	-0.020	-0.273***	0.011	-0.003
High performance	(0.187)	(0.144)	(0.097)	(0.049)	(0.022)
Treatment × Post			0.066***		0.056**
			(0.024)		(0.026)
Treatment × Post ×			0.123*		0.023
High performance			(0.066)		(0.028)
Obs.	2161	2246	9376	2246	8654
Dep. variable mean for control group	0.815	0.790	0.776	0.790	0.764
Pseudo R-sq.	0.24	0.22	0.13	0.22	0.33

Notes: Columns 1-3 include interaction terms with an indicator for above median share of 7th graders in the class who abstained at baseline. The median corresponds to 63% of the class abstaining while they were in 7th grade.Columns 4-5 interact treatment with above median share of 8th graders in the class who abstained at 3 months (while in 7th grade). The median corresponds to 75% of the class. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.10: Program effect on tobacco abstinence by students' risk preference

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Treatment	0.086***	0.069***	-0.019
	(0.026)	(0.026)	(0.017)
Certainty	0.001	0.033	0.010
	(0.028)	(0.023)	(0.017)
Treatment ×	-0.033	-0.017	0.032
Certainty	(0.037)	(0.034)	(0.023)
Treatment × Post			0.075***
			(0.027)
Treatment × Post ×			-0.022
Certainty			(0.027)
Obs.	2156	2245	9350
Dep. variable mean for control group	0.817	0.790	0.778
Pseudo R-sq.	0.24	0.22	0.13

Notes: District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table B.11: Program effect on tobacco abstinence by students' time preference

	(1)	(2)	(3)
	3-month	8-month	Diff-in-diff
Panel A. Impatient			
Treatment	0.074***	0.054**	-0.008
	(0.021)	(0.024)	(0.014)
Impatient	-0.014	-0.025	0.032
	(0.032)	(0.029)	(0.033)
Treatment ×	0.007	0.047	-0.012
Impatient	(0.046)	(0.048)	(0.041)
Treatment × Post			0.071***
			(0.024)
Treatment × Post ×			-0.021
Impatient			(0.042)
Obs.	2156	2245	9350
Dep. var. mean for control group		0.817	
R-sq.	0.24	0.22	0.13
Panel B. Time inconsistent			
Treatment	0.070**	0.047	0.010
	(0.030)	(0.033)	(0.021)
Time inconsistent	0.004	-0.018	-0.000
	(0.022)	(0.022)	(0.015)
Treatment ×	0.007	0.019	-0.029
Time inconsistent	(0.033)	(0.036)	(0.022)
Treatment × Post			0.044
			(0.033)
Treatment × Post ×			0.036
Time inconsistent			(0.032)
Obs.	2156	2245	9350
Dep. var. mean for control group	0.817	0.790	0.778
R-sq.	0.24	0.22	0.13

Notes: District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table B.12: Heterogeneous treatment effects by the last time a student smoked

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	>2 weeks			> 1 month			Ever smoked		
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Treatment	0.087**	0.095**	-0.374***	0.080**	0.071**	-0.063***	0.081*	0.131***	-0.002
	(0.038)	(0.038)	(0.068)	(0.034)	(0.033)	(0.018)	(0.044)	(0.049)	(0.030)
Last smoked	-0.017	-0.045	0.436***	-0.008	-0.013	0.116***	-0.009	-0.075	0.013
	(0.039)	(0.039)	(0.072)	(0.039)	(0.031)	(0.025)	(0.044)	(0.051)	(0.027)
Treatment ×	0.042	0.046	0.043**	0.100***	0.123***	0.099***	-0.142***	-0.180***	-0.136***
Last smoked	(0.036)	(0.038)	(0.021)	(0.032)	(0.027)	(0.017)	(0.035)	(0.037)	(0.021)
Treatment × Post			0.504***			0.143***			0.037
			(0.073)			(0.026)			(0.032)
Treatment × Post ×			-0.560***			-0.185***			-0.015
Last smoked			(0.076)			(0.026)			(0.029)
Obs.	2157	2246	8645	2157	2246	8645	2157	2246	8645
Dep. var. mean for control	0.816	0.790	0.764	0.816	0.790	0.764	0.816	0.790	0.764
Pseudo R-sq.	0.24	0.22	0.34	0.25	0.23	0.34	0.26	0.26	0.34

Notes: District fixed effects included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.13: Program effect on tobacco abstinence by students' own smoking intentions

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Treatment	0.076***	0.037	-0.034*
	(0.028)	(0.029)	(0.020)
High smoking probability	-0.111***	-0.121***	-0.212***
	(0.024)	(0.019)	(0.017)
Treatment ×	-0.000	0.042	0.044*
High smoking probability	(0.033)	(0.029)	(0.023)
Treatment × Post			0.058*
			(0.031)
Treatment × Post ×			0.012
High smoking probability			(0.026)
Obs.	2157	2246	9353
Dep. var. mean for control	0.816	0.790	0.777
R-sq.	0.26	0.24	0.19

Notes: High smoking probability takes the value one if the student believed he would smoke with a non-zero probability within three months at baseline. This corresponds to the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table B.14: Program effect on tobacco abstinence by parental smoking status

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Treatment	0.105***	0.055**	-0.005
	(0.024)	(0.026)	(0.017)
Parent smoked	-0.028	-0.071***	-0.072***
	(0.018)	(0.021)	(0.016)
Treatment ×	-0.058**	0.011	-0.008
Parent smoked	(0.028)	(0.032)	(0.022)
Treatment × Post			0.082***
			(0.024)
Treatment × Post ×			-0.026
Parent smoked			(0.022)
Obs.	2161	2246	9378
Dep. var. mean for control	0.815	0.790	0.776
R-sq.	0.24	0.22	0.13

Notes: District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

Table B.15: Program effect on tobacco abstinence by parental characteristics

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Panel A. Maternal education			
Treatment	0.075***	0.052**	-0.016
	(0.021)	(0.025)	(0.015)
Low maternal education	-0.013	-0.060**	-0.046
	(0.030)	(0.025)	(0.029)
Treatment ×	-0.002	0.065*	0.042
Low maternal education	(0.038)	(0.039)	(0.034)
Treatment × Post			0.067***
			(0.024)
Treatment × Post ×			0.010
Low maternal education			(0.028)
Obs.	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776
Pseudo R-sq.	0.24	0.22	0.13
Panel B. Paternal education			
Treatment	0.072***	0.057**	-0.011
	(0.020)	(0.024)	(0.015)
Low paternal education	0.012	0.026	0.018
	(0.038)	(0.028)	(0.020)
Treatment ×	0.017	0.030	0.007
Low paternal education	(0.048)	(0.044)	(0.033)
Treatment × Post			0.064***
			(0.024)
Treatment × Post ×			0.027
Low paternal education			(0.031)
Obs.	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776
Pseudo R-sq.	0.24	0.22	0.13
Panel C. Low occupational status			
Treatment	0.064**	0.027	0.061**
	(0.029)	(0.033)	(0.027)
Low occupation	-0.029	-0.038	-0.036*
	(0.024)	(0.024)	(0.019)
Treatment ×	0.014	0.047	-0.070**
Low occupation	(0.032)	(0.038)	(0.030)
Treatment × Post			-0.037
			(0.028)
Treatment × Post ×			0.076***
Low occupation			(0.024)
Obs.	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776
Pseudo R-sq.	0.24	0.22	0.13

Notes: Low education takes the value one if the parent has primary education or lower. Low occupation includes agriculture and blue collar workers. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.16: Program effect on tobacco abstinence by smoking among teachers in the school

	(1)	(2)	(3)
	3-month	8-month	Diff-in-diff
Treatment	0.081***	0.055*	0.013
	(0.022)	(0.030)	(0.018)
No teachers smoke	-0.021	-0.016	-0.011
	(0.033)	(0.040)	(0.023)
Treatment ×	-0.056	0.015	-0.037
No teachers smoke	(0.043)	(0.055)	(0.033)
Treatment × Post			0.075***
			(0.024)
Treatment × Post ×			-0.043
No teachers smoke			(0.049)
Obs.	2161	2246	7182
R-sq.	0.24	0.22	0.14

Notes: 30% of schools reported no teachers smoking at baseline. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.17: Program effect on tobacco abstinence by school's academic performance

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Treatment	0.044*	0.005	-0.030*
	(0.025)	(0.027)	(0.017)
High performance	0.023	-0.027	0.006
	(0.031)	(0.033)	(0.020)
Treatment ×	0.070**	0.131***	0.045*
High performance	(0.036)	(0.050)	(0.026)
Treatment × Post			0.053**
			(0.027)
Treatment × Post ×			0.052
High performance			(0.033)
Obs.	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776
R-sq.	0.24	0.23	0.13

Notes: High performance corresponds to above median mathematics test score. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.18: Program effect on tobacco abstinence by school location

	(1)	(2)	(3)	(4)	(5)	(6)
		Urban			Near treatment school	
	3-month	8-month	Diff-in-diff	3-month	8-month	Diff-in-diff
Treatment	0.078***	0.072***	-0.016	0.051*	0.073**	-0.010
	(0.023)	(0.028)	(0.016)	(0.028)	(0.031)	(0.017)
Variable	0.024	0.054	0.008	-0.014	-0.015	-0.017
	(0.034)	(0.039)	(0.023)	(0.031)	(0.037)	(0.020)
Treatment ×	-0.011	-0.037	0.030	0.051	-0.024	-0.000
Variable	(0.038)	(0.058)	(0.030)	(0.037)	(0.049)	(0.026)
Treatment × Post			0.081***			0.059**
			(0.024)			(0.024)
Treatment × Post ×			-0.057			0.019
			(0.044)			(0.034)
Obs.	2161	2246	9378	2161	2246	9378
Dep. var. mean for control group	0.815	0.790	0.776	0.815	0.790	0.776
R-sq.	0.24	0.22	0.13	0.24	0.22	0.13

Notes: Near takes the value one if a school is less than 1.5 kilometers away from the nearest treatment school. This corresponds to the median. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors are clustered at the school level.

Table B.19: Alternative outcomes

	(1) 3-month	(2) 8-month	(3) Diff-in-diff
Panel A. Cotinine test only			
Treatment	0.043**	0.039*	0.009
	(0.018)	(0.021)	(0.011)
Treatment × Post			0.026
			(0.019)
Obs.	2161	2246	6757
Dep. var. mean for control group		0.857	
Pseudo R-sq.	0.25	0.24	0.40
Panel B. Friend report supplement			
Treatment	0.059***	0.052**	0.008
	(0.022)	(0.025)	(0.012)
Treatment × Post			0.041*
			(0.021)
Obs.	2268	2341	6959
Dep. var. mean for control group	0.772	0.746	0.774
Pseudo R-sq.	0.21	0.20	0.35
Panel C. Including missing test			
Treatment	0.067***	0.059**	0.009
	(0.019)	(0.023)	(0.011)
Treatment × Post			0.049**
			(0.021)
Obs.	2268	2341	6959
Dep. var. mean for control group	0.824	0.799	0.776
Pseudo R-sq.	0.21	0.20	0.36

Notes: Panel C assumes a student who did not take the cotinine test in subsequent waves did not smoke if he abstained at baseline. District fixed effects and baseline outcome included. Fully adjusted includes: school's average math test score, the number of male teachers who smoke, household size, household land ownership, skilled occupation of the head of the household, parental smoking status. Standard errors clustered at the school level.

The Persistent Effects of Early-Life Exposure to Air Pollution: Evidence from the Indonesian Forest Fires

Maria Rosales-Rueda and Margaret Triyana*

February 20, 2018

Abstract

We analyze the effects of early-life exposure to air pollution in a developing country on children's long-term human capital outcomes. We exploit the geographical variation of the 1997 Indonesian forest fires and cohort variation in exposure as a natural experiment. Children exposed to the fires are shorter on average three years post-exposure and have lower lung capacity 10 years post-exposure, but only children who were exposed in-utero continue to exhibit shorter stature at 10 and 17 years post-exposure. We find suggestive evidence that these persistent effects may be due to expectant mothers experiencing poorer respiratory health during the fires.

Keywords: natural disasters, air pollution, long-term effects, fetal origins, Indonesia

*Maria Rosales-Rueda is a Visiting Assistant Professor at the Department of Economics and the Center for Health and Wellbeing at Princeton University (mrosales@uchicago.edu). Margaret Triyana is a Visiting Assistant Professor of Economics at the Keough School of Global Affairs at the University of Notre Dame (mtriyana@uchicago.edu). The data used in this article can be obtained beginning six months after publication through three years hence from Maria Rosales-Rueda (mrosales@uchicago.edu). We are grateful to our beloved advisor, Bob Lalonde, for his continued guidance and support. We thank three anonymous referees for detailed and thoughtful comments that greatly improved the paper. We also thank Marinho Bertanha, Janet Currie, Kirk Doran, Bill Evans, Nilesh Fernando, Lakshmi Iyer, Joe Kaboski, Julien Labonne, John Strauss, Anja Tolonen, Tom Vogl, and seminar participants at the University of Notre Dame, Princeton University, the National University of Singapore, the University of Houston, London School of Economics, Monash University, the University of Bath, the University of Adelaide, the University of Delaware, and the World Bank for valuable suggestions. Stephanie Mastorakos and Katherine Fugate provided excellent research assistance. This research was financially supported by a Nanyang Technological University Start-up Grant.

doi:10.3368/jhr.54.4.0117.8497R1

1 Introduction

Burning organic matter, which includes biomass burning, open fires, agricultural and forest fires, is a major source of air pollution, accounting for 85% of airborne particulate pollution (Landrigan et al., 2017). One of the main sources of outdoor air pollution is slash and burn agriculture– an activity that is estimated to involve about 30 million people globally (Dixon, Gibbon and Gulliver, 2001). Weather patterns like droughts can exacerbate the fires by delaying the rain necessary to extinguish them. With climate change, weather phenomena are expected to become more unpredictable, making wild fires occur more frequently and more difficult to control (United Nations Framework Convention on Climate Change, 2007). Moreover, people living in low and middle income countries are disproportionately affected by the burden of air pollution, with 87% of its premature deaths occurring in these countries.[1] A better understanding of how these environmental shocks affect human capital formation can shed light on the latent burden of air pollution, especially in developing countries. This paper analyzes the effects of early-life exposure to air pollution on long-term human capital outcomes in Indonesia.

We focus on the long-term effects of early-life exposure to air pollution because children face a high risk of contracting pollution-related diseases as exposure to pollutants during sensitive periods in-utero and early childhood can result in lasting damages (Landrigan et al., 2017). A large literature has documented the adverse short-term effects of exposure to air pollution early in life on child health (Chay and Greenstone, 2003; Arceo, Hanna and Oliva, 2016; Foster, Gutierrez and Kumar, 2009; Rangel and Vogl, 2016). These findings have motivated the study of the long-term effects of such exposure, and a growing body of evidence has shown that the documented short-term impacts translate into poor adult well-being (Chay and Greenstone, 2003; Kajekar, 2007; Currie et al., 2014; Currie and Vogl, 2013 for a review). However, the majority of these studies have examined the long-term impact of early-life exposure to air pollution in high income countries, and little is known about the magnitudes of long-term effects in developing countries. Our paper fills this gap by examining the persistent effects of early-life exposure to air pollution from a relevant source in low and middle-income countries: burning organic matter. Due to limited resources for mitigation in developing countries, the long-term effects of air pollution may be more severe for children in low income countries relative to children in high income countries (Arceo, Hanna and Oliva, 2016; Hanna and Oliva, 2016).

Previous studies have relied on natural experiments such as radioactive emissions, wild

[1] http://www.who.int/mediacentre/factsheets/fs313/en/
Last accessed December 28, 2017.

fires, changes in government regulations, economic crises, and plant closings to estimate the causal effects of early-life exposure to air pollution on short and long-term outcomes (Almond, Edlund and Palme, 2009; Bharadwaj, Eberhard and Neilson, 2010; Currie and Vogl, 2013; Currie et al., 2014). However, assessing the long-term impacts of exposure to air pollution is challenging because some events that might change particulate levels may also generate longer-term changes at the community level (e.g. the decline of the manufacturing industry or changes in the Clean Air Act). Consequently, these changes would confound the impact of the change in pollution with general-equilibrium effects. Studies that analyze the long-term effects of early-life exposure to air pollution have relied on short, acute episodes like the Chernobyl accident that affected a well-defined population during sensitive periods of human capital formation (Almond, Edlund and Palme, 2009). Similarly, we exploit exposure to the Indonesian fires in 1997, which burned more intensely due to the El Niño weather phenomenon that caused extremely dry conditions and delayed the monsoon rain in Indonesia. The unprecedented intensity, sudden timing, and additional duration of the fires provide a source of exogenous variation in cohort exposure to air pollution that we use as a natural experiment (Frankenberg, McKee and Thomas, 2005; Jayachandran, 2009).

The 1997 Indonesian forest fires occurred between August and November and primarily affected the islands of Sumatra and Kalimantan. We exploit the geographic and cohort variation to estimate the effects of early-life exposure to air pollution. In our estimation, we combine data from NASA's Earth Probe Total Ozone Mapping Spectrometer (TOMS) and the Indonesian Family Life Survey (IFLS), which includes communities in affected and unaffected areas and provides long-term outcomes. More specifically, data from TOMS provide daily measures of air pollution during the fires, while data from the IFLS were collected during the fires, followed by 3, 10, and 17 years after the 1997 fires. The rich longitudinal data from the IFLS includes children's health status, anthropometric measures, lung capacity, and cognitive tests. We use children's month, year, and place of birth to identify children who were exposed to the fires in-utero, in their first year of life, and/or their second year of life.

These Indonesian forest fires have been studied in previous literature to mainly examine the short-term effects of air pollution in Indonesia and its neighboring countries (Emmanuel, 2000; Frankenberg, McKee and Thomas, 2005; Jayachandran, 2009; Sastry, 2002). Frankenberg, McKee and Thomas (2005) study the immediate effects of the fires on adult health, while Jayachandran (2009) examines the immediate consequences of exposure to the fires on early-life mortality in Indonesia. We focus on estimating the longer-term effects of exposure to these fires on surviving children's human capital at 3, 10, and 17 years post-exposure and the potential mechanisms behind the persistent effects. Although we rely on a well-known

extreme pollution event for causal identification, events of similar magnitude are likely to occur every 15-20 years (McPhaden, 2015); in fact, the intensity of the 2015 forest fires in Indonesia is comparable to the 1997 event. Moreover, forest fires are episodic events whose frequency may increase due to climate change (United Nations Framework Convention on Climate Change, 2007), and even high income countries are not exempt from their occurrence, as California recently experienced wild fires that were responsible for one of the worst air quality episodes in record. In addition, the increase in the levels of particulate matter generated by the 1997 forest fires is similar to the hazardous haze from agricultural burning that affects cities like New Delhi. The haze from forest fires is also comparable to the chronic exposure to indoor air pollution generated by the use of biomass fuels (Hanna, Duflo and Greenstone, 2016).

When examining the long-term effects of early-life adversities, there may be competing effects between selective mortality and scarring effects associated with air pollution. We may observe the paradox of mortality selection: under high selective mortality, long-term scarring may not be observed; with low mortality, long-term scarring is observed. The 1997 fires in Indonesia were associated with lower cohort size, but the magnitude of the effect is relatively small compared to other large scale natural disasters (Almond and Currie, 2011; Jayachandran, 2009). With relatively low mortality, scarring may be observed in the long term, so these surviving children may have poor human capital outcomes in the longer term as a result of exposure to the fires. Other studies, mostly from high-income countries, have found that exposure to pollution during the first year of life is associated with lasting negative effects on adult outcomes, including cognitive function, education, and labor market outcomes (Bharadwaj et al., 2017; Isen, Rossin-Slater and Walker, 2014; Peet, 2015; Sanders, 2012; Almond et al., 2009; Black et al., 2013; Currie and Vogl, 2013 for a review). We find that surviving children's health is also adversely affected 3, 10, and 17 years post exposure in a middle income country.

Our results suggest that the health effects persist, as evidenced by exposed children's lower height. We find that children exposed to air pollution in-utero have lower height-for-age z-scores three years post exposure, and these same children are also on average 0.2 standard deviations shorter later in life at 10 and 17 years post-exposure. Furthermore, children who were exposed in-utero or in their first 2 years of life have 5 to 9 percent lower lung capacity 10 years post-exposure (note that this effect is no longer significant 17 years post-exposure). However, there appears to be no significant effects on children's cognitive function. We find limited heterogeneous treatment effects by maternal education or urban residence, suggesting that the health effects of exposure to air pollution are not mitigated by socio-economic status. Regarding potential mechanisms underlying these effects, we find

suggestive evidence that contemporaneous effects on maternal and infant health may explain the persistent health impacts. During and immediately after the fires, expectant mothers experienced a deterioration in their lung capacity and children under the age of two suffered breathing problems. We also show evidence that selective migration or fertility responses are not likely to confound our results. Examining the impact of early-life exposure to pollution in low and middle income countries is of particular interest since pollution is an important policy challenge in such countries. The lasting consequences of air pollution on children's human capital may not be apparent immediately, but hidden in future damages.

The remainder of the paper is organized as follows. Section 2 presents the background on the 1997 fires in Indonesia. Section 3 describes the data and estimation strategy. Section 4 presents the results, and section 5 provides a short discussion and concludes.

2 Background

2.1 Fires in Indonesia

In many low and middle income families, farmers, especially small scale ones, have traditionally used controlled burning as a method of land clearing. Land clearing happens before planting and, in Indonesia, takes place during the dry season, typically between July and October. This method of clearing is cheaper than its alternative, which involves heavy machinery and requires higher upfront capital investments. Historically, land clearing was limited to domestic settings, so the scope and effect of the fires were small. However, over time, the scale of the fires increased as timber and palm oil plantations developed. These industries created more flammable debris and required larger areas of land to be cleared. In addition, peat land in parts of Sumatra and Kalimantan makes it difficult to contain fires. Because peat contains a mixture of decaying organic matter, when exposed to drought conditions, it subsequently dries and combusts easily, thereby causing fires to spread both quickly and usually underground. These conditions make it difficult to extinguish these fires without the help of heavy rainfall. In 1995, Indonesia banned the practice of use of fire to clear land. Since then, however, farmers have ignored this ban, and the magnitude of the fires has continued to increase (Sastry, 2002).

Between August 1997 and November 1997, Indonesia experienced the most intense and long lasting series of forest fires in its history. These fires resulted from both annual land clearing by fire and the El Niño-Southern Oscillation weather phenomenon. El Niño is associated with unusual weather patterns around the world and causes extremely dry conditions in Indonesia, resulting in droughts and delaying Indonesia's monsoonal weather pattern.

Consequently, once the Indonesian farmers began land clearing by burning, the drought conditions caused by El Niño exacerbated the fires, causing them to rage out of control and forcing Indonesia's then president Suharto to declare a state of emergency. The islands of Sumatra and Kalimantan were most affected by these fires, which destroyed 2 to 3 percent of Indonesia's land area and produced thick haze that spread to neighboring Southeast Asian countries such as Brunei, Malaysia, Singapore, Thailand, and Vietnam (Fullerton, Bruce and Gordon, 2008; Sastry, 2002). The estimated cost of these fires ranged from $2 to 3 billion in lost productivity and tourism, which was about 1 to 1.4 percent of GDP in 1997 (Tacconi, 2003). The El Niño weather phenomenon happened again in 2015, causing about 100,000 deaths in the region as the fires once again burned intensely in Indonesia, producing thick haze that also spread to neighboring countries (Koplitz et al., 2016).

The 1997 Indonesian fires resulted in increased levels of pollution as measured by total suspended particulate matter (TSP) and PM_{10}. TSP denotes the number of airborne particles or aerosols that are less than 100 micrometers, while PM_{10} designates smaller particles with a diameter of 10 micrometers or less. The smaller PM particles are able to reach the lower regions of the respiratory tract, and, consequently, can be responsible for adverse health effects such as inhibited breathing and respiratory system function, damage to lung tissue, cancer, and premature death. The elderly, children, and people with chronic lung disease, influenza, or asthma, are especially sensitive to the effects of particulate matter. The main gases produced by these fires are carbon dioxide (CO_2), carbon monoxide (CO), methane (CH_4), nitrogen oxides (NO and NO_2), and ammonia. The peat land causes a higher level of methane to be emitted compared to fires in areas without peat. The PM_{10} standard set by the US Environmental Protection Agency (EPA) is an average level of no more than 150 $\mu g/m^3$ within a 24-hour period in a location more than once per year. Earlier standard for the US used a TSP level of 260 $\mu g/m^3$ within a 24-hour period in a location more than once per year.[2] According to the WHO guidelines for air quality, the PM_{10} 24-hour average standard is 50 $\mu g/m^3$. During the 1997 Indonesia fires, the PM_{10} levels reached the hazardous range for the most affected provinces, surpassing 2000 $\mu g/m^3$ in September, while TSP levels ranged between 150 and 4000 $\mu g/m^3$ (Heil and Goldammer, 2001; Ostermann and Brauer, 2001). The level of pollution associated with these fires is comparable to levels of pollution associated with the chronic exposure to indoor air pollution generated by the use of biomass fuels (Hanna, Duflo and Greenstone, 2016).

These pollution measures have been shown to be correlated with the aerosol index recorded by NASA's Earth Probe Total Ozone Mapping Spectrometer (TOMS) (Frankenberg, McKee and Thomas, 2005; Torres et al., 2002). The TOMS aerosol measurements

[2]https://www3.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html Last accessed January 12, 2018

were taken daily from July 1996 to December 2005 in one degree latitude and 1.25 degree longitude grids. The index ranges from to 0 to 6 in Indonesia, where values of less than 0.1 correspond to clear skies and a value of 4 or greater corresponds to low visibility that makes it difficult to see the mid-day sun. During the fires, when the aerosol index recorded a value of close to 6 in the most affected areas, PM_{10} levels in those areas were above 3000 $\mu g/m^3$ (Frankenberg, McKee and Thomas, 2005).

The source of variation in exposure to air pollution exploited in this paper is similar to the one used by Frankenberg, McKee and Thomas (2005), Jayachandran (2009), Banerjee (2016), and Sheng (2017). Jayachandran (2009) studies fetal, infant, and child mortality using data from the 2000 Indonesian census as well as pollution measures from the TOMS aerosol index. She finds that higher levels of air pollution due to the 1997 fires led to a decline of 1.2 percent in cohort size, especially due to exposure during the third trimester in-utero. She also finds that the effects are twice as large in poorer districts. These findings are consistent with previous studies that have found that exposure to pollution in-utero affects children's health at birth and increases the likelihood of neonatal and infant mortality (Chay and Greenstone, 2003; Arceo, Hanna and Oliva, 2016; Foster, Gutierrez and Kumar, 2009; Currie et al., 2014; Rangel and Vogl, 2016 Currie and Vogl, 2013 for a review). Additionally, Frankenberg, McKee and Thomas (2005) study the consequences of the fires on adult health using data from the IFLS and pollution data from TOMS. They find that the fires had a negative impact on contemporaneous self-reported physical health, as measured by difficulty carrying a heavy load, respiratory problems, and general health status. We complement this literature by examining the effects of early-life exposure to air pollution on surviving children's human capital outcomes in the longer term using additional waves of the IFLS and pollution measure from TOMS.

2.2 Air pollution and child development

The prenatal and early childhood periods are sensitive stages of child development, and negative shocks such as pollution during these years may predispose children to poor outcomes later in life (Barker, 1995; Currie and Vogl, 2013). For example, negative shocks during these early years can impair skill formation, which in turn can result in long-term consequences since past human capital investments influence both future skills and investments (Cunha and Heckman, 2007).

In-utero exposure to air pollution has been shown to adversely affect morbidity and mortality since a child's lung development commences in-utero (Currie, Neidell and Schmieder, 2009; Jayachandran, 2009). Furthermore, children remain vulnerable to the effects of air

pollution as their lung development continues postnatally. For example, the elevated risk of altering the development of lung function when children are exposed to viral infections during infancy suggests that early-life exposure to air pollution can be equally damaging (Kajekar, 2007). While studies have shown the effects of early-life exposure to air pollution on neonatal and infant mortality, as well as children's lung function (Currie and Vogl, 2013), we focus on children's long-term lung function and other health outcomes, conditional on surviving the initial exposure to air pollution.

One health outcome that we examine is children's height both in the short and long-term. Height is a widely used proxy of health status that captures the cumulative effects of health shocks and previous research has shown its link to other dimensions of well-being like labor market outcomes (Alderman, Hoddinott and Kinsey, 2006; Case and Paxson, 2008; Currie and Vogl, 2013; Magnusson, Rasmussen and Gyllensten, 2006). Studies have also shown correlational evidence on the association between exposure to air pollution in early life and children's physical growth (Bobak, Richards and Wadsworth, 2004; Ghosh et al., 2011; Kyu, Georgiades and Boyle, 2009; Mishra and Retherford, 2007; Rona, 1981; Vik et al., 1996).[3] In-utero exposure to air pollution has been linked to intrauterine growth restriction (birth weight under the 10^{th} percentile by gestational age and sex) and shorter birth length (Ghosh et al., 2011). Children who were exposed to air pollution in-utero appear to catch up in weight by age 5, but not in height (Vik et al., 1996).

While the precise channels that link air pollution and children's height is uncertain, two hypothesized mechanisms are through carbon monoxide (Mishra and Retherford, 2007) and/or respiratory infections. Under the first hypothesis, carbon monoxide binds with hemoglobin, which is necessary for transporting oxygen to body tissues. As a result of the binding, carboxyhemoglobin forms, and this in turn reduces the quantity of hemoglobin in the blood, thereby causing anemia and subsequently, lower height. Exposed children's physical growth may also be impeded through lower birth weight (Rona, 1981). Under the second hypothesis, the medical literature argues that inhalation of small particles in sensitive periods of growth can retard body size by diverting the body's energy to fighting respiratory infections rather than developing bones and muscles ((Heinrich and Slama, 2007; Rückerl et al., 2011). Empirical evidence from historical data on coal-fired industrialization in the UK suggests that the association between height and exposure to air pollution in-utero and postnatally comes from respiratory infections (Bailey, Hatton and Inwood, 2016). In this paper, we explore the causal link and potential channels between acute exposure to air pollution in early life and children's height at 3, 10, and 17 years post exposure.

[3]There is also causal evidence on the relationship between fetal exposure to air pollution and delayed skeletal development from animal studies (Ungvary and Tatrai, 1985).

We also explore the effects of pollution on children's cognitive function. While the precise mechanism that links air pollution to cognitive development is currently unknown, the hypothesized channel is that carbon monoxide crosses the placental barrier, which subsequently negatively affects the cardiovascular and respiratory functions that influence cognitive outcomes (Greingor et al., 2001). Exposure to air pollution, especially carbon monoxide, in the first year of life is associated with poor infant and child health (Currie, Neidell and Schmieder, 2009; Mortimer et al., 2008; Neidell, 2004; Plopper and Fanucchi, 2000), including lower lower IQ scores (Bharadwaj et al., 2017) and lower educational attainment (Sanders, 2012). While the scope of this study and the data we utilize take into account cognitive function, it is possible that investments or later interventions could mitigate the initial effects of negative health shocks (Almond and Mazumder, 2013). It is therefore an empirical question whether children's cognitive function would be affected 10 and 17 years post exposure.

3 Data and Empirical Strategy

3.1 Data

This paper combines data of children and households from the Indonesian Family Life Survey (IFLS) and pollution data from NASA's Earth Probe Total Ozone Mapping Spectrometer (TOMS).

3.1.1 Indonesian Family Life Survey

The IFLS is a longitudinal household survey that is representative of approximately 83 percent of the Indonesian population in 1993. The first wave of the survey (IFLS1) was conducted in 1993, the second wave (IFLS2) in 1997, the third wave (IFLS3) in 2000, the fourth wave (IFLS4) in 2007, and the fifth wave (IFLS5) in 2014.

Indonesia is administratively divided into provinces, districts (regencies or cities), subdistricts, and villages in rural areas or townships in urban areas. The IFLS over sampled urban areas and rural areas outside of the main island of Java. The survey includes provinces affected by the 1997 fires on the islands of Sumatra and Kalimantan as well as other provinces that were not affected. IFLS1 included 7,224 households residing in 13 of Indonesia's 26 provinces in 1993. These households resided in approximately 200 districts, which corresponded to 321 enumeration areas in 312 communities. A community is defined as a village in rural areas and a township in urban areas. Subsequent waves of the survey sought to re-interview all households in IFLS1, as well as any households that had split-off.

The IFLS contains rich information on individual, household, and community characteristics. Individual characteristics include age, education, marital status, employment, as well as complete pregnancy history for women between the ages of 15 and 49. The following anthropometric measures were also taken: height, weight, blood pressure, and lung capacity. Household characteristics include household size, household expenditure, and asset ownership.

The timing of the IFLS surveys allows us to study the effects of early-life exposure to the 1997 fires at different points in children's life-cycle. During the second survey (IFLS2), almost 90 percent of the interviews were conducted between June and December 1997, 74 percent of which were conducted while the fires were burning between June and November 1997. Therefore, IFLS2 allows us to analyze the immediate effects of the fires on children's and mothers' contemporaneous outcomes. IFLS3, collected in 2000, makes it possible for us to analyze the short-term effects of the fires, 3 years post-exposure. IFLS4 and IFLS5, conducted in 2007 and in 2014 respectively, enable us to examine longer-term outcomes including height, cognitive test scores and lung capacity.

Outcomes In this study, children's outcomes come from the individual surveys, the anthropometric, and cognitive measures. Since stunting captures the cumulative effect of negative health shocks on children's physical growth, the short-term outcome we examine is children's height-for-age in 2000, when the children in our sample are under 5 years of age. The indicator for stunting takes the value 1 if a child's height-for-age is more than 2 standard deviations below the standardized mean. Trained enumerators measured children's height and weight in each wave of the survey. We used the WHO standard to generate z-scores for children's height-for-age as a measure of children's health status.

Our longer-term outcomes are children's height, lung capacity, and cognitive development. We use children's height in 2007 and 2014 as our first longer-term health outcome because height has been commonly used to capture the cumulative effects of health shocks and is a marker of long-term adult well-being (Case and Paxson, 2008; Currie and Vogl, 2013).

As our second longer-term outcome, we use lung capacity, which measures peak respiratory flow, which indicates a person's maximum speed of expiration/exhalation in liters per minute. This measure shows whether there are any obstructions in the respiratory pathways. For this study, we use the IFLS' measures of lung capacity taken from respondents older than 9 years of age.[4] For the children in our sample, lung capacity is only available in 2007 and 2014. In each survey, three measures were taken for each respondent. Following

[4]In 2007 (IFLS4), this measure is not available for children born later than 1999.

the literature, we use the maximum of the three measurements (Silwal and McKay, 2014).

The IFLS administers cognitive tests on children and adult respondents between the ages of 7 and 24 years. In this study, we use test scores from 2007 and 2014, when the children are old enough to analyze their longer-term outcome. Cognitive function is measured using 2 outcomes: (1) the Raven progressive matrices, which were designed to assess abstract reasoning (such as spatial and problem-solving skills), and (2) the mathematics test, which determines a respondent's numeracy skill. Children between 7 and 14 years old were given 12 questions from the Raven's test and 5 mathematics questions. Each child's score is the percentage of correct answers to the Raven and mathematics tests.[5] We combine these scores as our cognitive outcome.[6]

Potential mechanisms We also explore potential mechanisms behind the short and longer-term effects of exposure to the fires. In the 1997 survey (IFLS2), the IFLS collected information on respiratory problems. An indicator for respiratory problems takes the value 1 if the mother or primary care giver reported the child (aged 3 years of younger) had experienced any difficulty breathing or any coughing in the month prior to the second wave of the survey. Since a sizable proportion of the surveys were conducted during or immediately after the fires, this indicator captures the immediate effect of the fires on young children's health, which is one potential channel for children's weaker respiratory function in the longer-term. We also examine wasting, defined as below two standard deviations in children's weight-for-height as an indicator of acute malnutrition, which is potentially caused by infectious respiratory diseases. In addition, we explore the effects of the shock on the child's birth weight since it is considered a measure that summarizes prenatal investments in-utero (Bharadwaj, Eberhard and Neilson, 2010; Rosenzweig and Zhang, 2009).

Other potential mechanisms underlying the in-utero effects come from the impacts of the fires on expectant mothers' health. Therefore, we analyze mother's respiratory systems based on their lung capacity. We also study whether pregnant women's nutrition may have been disrupted by looking at mother's body mass index (BMI), which is defined as a person's weight in kilograms divided by her height in meters squared. Lastly, we also examine the effects on investment responses through breastfeeding and fertility decisions.

[5]The cognitive measure was restricted to children who completed both the cognitive and mathematics components.

[6]We separate the tests into the Raven and mathematics and find qualitatively similar estimates.

3.1.2 Pollution data: NASA's Earth Probe Total Ozone Mapping Spectrometer (TOMS)

Our data on pollution come from TOMS (version 8). The daily aerosol index from TOMS has been shown to be correlated with levels of Total Suspended particles (TSP) (Torres et al., 2002). TOMS measures daily aerosol in one degree latitude[7] and 1.25 degree longitude grids. Adjacent grid points are approximately 175 kilometers apart. For this study, we use the available aerosol index from July 1996 to December 2005. The index ranges from 0 to almost 6 during the fires. We use the monthly average of the aerosol index for each grid and match these grids to the nearest IFLS district centroid.[8] There are 65 grids that correspond to the 205 IFLS districts in our analyzed sample. The land area of Indonesian districts range anywhere from under 20 kilometers squared to 45,000 kilometers squared. A range of 1 to 9 IFLS districts were matched, depending on size, to one TOMS grid, with an average of 3 districts per TOMS grid.

We define a district as exposed to the 1997 forest fires if the aerosol index exceeded 1.5 for more than 3 days in any month between August 1997 and November 1997 (Frankenberg, McKee and Thomas, 2005). An aerosol index of 1.5 corresponds to a value of around 500 $\mu g/m^3$ in TSP.[9] To link the aerosol index to the occurrence of forest fires in this period, Figure 1 displays the co-occurrence of our pollution exposure variable and burnt area in September 1997 using data from the fourth-generation Global Fire Emissions Database (GFED4) (Giglio, Randerson and Werf, 2013).[10] The dispersion of the aerosol index is also consistent with the general wind direction (Heil and Goldammer, 2001). More formally, as shown in Table A.1 in the appendix, there is a strong positive association between monthly burnt area and both the TOMS mean aerosol index and the number of days that the aerosol index is greater than 1.5 within a month.

3.1.3 Summary statistics

We restrict the sample to children born between 1995 and 2000. We then divide the sample into two groups: exposed children and non-exposed children. Exposed children

[7] Corresponding to 69 miles (111 kilometers)

[8] Indonesia experienced district proliferation over the years, so we use IFLS district centroids in 1997.

[9] This is almost double the US EPA threshold of unhealthy air quality (above 260 $\mu g/m^3$. See https://www3.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html Last accessed January 12, 2018). During the fires, an aerosol index of 3 corresponds to a TSP level of 1200 $\mu g/m^3$ and at the peak of the fires, the aerosol index reached 6, which corresponds to a TSP level of about 4000 $\mu g/m^3$. We estimated these values based on estimated TSP in affected areas in Sumatra and Kalimantan as reported by Heil (1998) and the aerosol index for the TOMS grid that includes these areas in September 1997.

[10] This data set contains the monthly burnt area from mid-1995 to the present at 0.25 degree spatial resolution. We match burned area in 1997 and 1998 to the pollution grids from TOMS.

include children who were exposed to an aerosol index of more than 1.5 for at least 3 days in a month when the child is in-utero, between 0 to 12 months, or 13 to 24 months. Non-exposed children include all other children born in areas affected by the fires or children who were born in areas not affected by the fires. All of these children are old enough for us to examine their longer-term outcomes.[11] Altogether, the sample includes around 3,700 children born between 1995 and 2000 and observed in 2014 (IFLS5).[12] Among these children, 10 percent were born in districts that experienced the 1997 fires.[13] On average, children who were affected by the fires in-utero were exposed to high pollution for about 15 days.

Table 1 presents our sample characteristics. Panel A shows children's outcomes in 2000, 3 years after the fires. Panels B and C present children's outcomes in 2007 and 2014, 10 and 17 years after the fires respectively. Children who were exposed to the fires were on average shorter for their age (height-for-age z-score of -1.5 for the non-exposed vs. -2.0), and a higher fraction of exposed children were stunted (38 percent for the non-exposed vs. 46 percent) in the survey in 2000 (IFLS3). The average height in 2007 (IFLS4) of exposed children is 133 centimeters, while those who were not exposed (who are generally younger than exposed children) are about 127 centimeters. The log lung capacity of children who were exposed and not exposed is similar at about 5.3, 10 years after the fires. For cognitive development, both groups correctly answered about 70 percent of the test questions. Seventeen years after the fires, children who were exposed and not exposed are about the same height, have the same lung capacity, and achieve the same test scores (in IFLS5).

3.2 Estimation strategy

In order to examine the short, medium and long-term effects of early-life exposure to pollution on human capital outcomes, we use a natural experiment design that exploits cohort variation in the timing of exposure (first difference) and geographic variation in the incidence and intensity of the 1997 forest fires (second difference). We employ a difference-in-differences approach to estimate the following model:

[11] Although migration can be an affected household's response to the fires, because the decision to migrate is endogenous, we include migrants in our analyzed sample. Approximately 4 percent of children born in both affected and non-affected districts moved between 1997 and 2000. When migrants are excluded, the estimates are qualitatively similar.

[12] The number of children varies by outcomes and waves.

[13] Table A.2 shows that the results are robust if we further restrict the sample to children whose height measurements were taken in 2000, 2007, and 2014. Since height measurements were taken in every IFLS wave, this restriction would allow us to consistently compare children's outcomes over time. The results are qualitatively similar.

$$y_{idmy} = \delta_0 inutero_{idmy} + \delta_1 age1_{idmy} + \delta_2 age2_{idmy} + \beta X_i + \tau_y + \mu_m + \rho_d + \epsilon_{idmy} \quad (1)$$

where y_{idmy} is the outcome of interest for individual i, born in district d at month m and year y. The variable $inutero_{idmy}$ captures exposure to the 1997 fires in-utero, $age1_{idmy}$ captures exposure between 0 to 12 months, and the variable $age2_{idmy}$ captures exposure between 13 and 24 months. Exposure to air pollution is calculated based on a child's month and year of birth and district of birth from the IFLS. We assume 9 months of gestation for all children in our sample.[14] In-utero exposure takes the value one if any month of the child's gestation period registered an aerosol index greater than 1.5 for more than three days during the 1997 forest fires in his/her district of birth. Similarly, exposure in the first (second) year of life takes the value one if any month of the child's first (second) year of life experienced an aerosol index greater than 1.5 for more than three days during the 1997 forest fires in his/her district of birth. Figure A.1 shows the average aerosol index early in life for children born between 1995 and 2005. Average exposure is similar across cohorts in affected and unaffected areas pre and post-1997. However, in 1997, children residing in affected districts were exposed to a higher level of pollution. For example, panel A shows that children who were in-utero during the fires (1998 birth cohort) were exposed to more pollution in-utero in the affected areas (panel A). Similarly, children who were born in 1997 and were between 0 and 12 months old in that year were exposed to higher pollution levels in the affected areas (panel B). Likewise, children who were between 13 and 24 months old in 1997 (those born in 1996) were also exposed to higher pollution in the affected areas. X_i is a vector of maternal and child characteristics that includes: male child, mother's age at delivery, and mother's education, urban residence, household size, household per capita expenditure (log), an asset index and a housing quality index.[15] We include year and month of birth fixed effects, τ_y and μ_m respectively, to take into account the common shocks that affect children born in the same month and year. We include district of birth fixed effects, ρ_d, to take into account time invariant district characteristics. Standard errors are clustered at the pollution exposure grid level.[16] The key coefficients of interest are δ_0, δ_1 and δ_2, which measure the effect of exposure to air pollution early in life on the outcome of interest.

[14] The average gestational age is 39 weeks and a standard deviation of 2.01. Using the available gestational age, we find that the fires do not significantly change gestational age (the estimated coefficient is 0.022, with a standard error of 0.099).

[15] The asset index is composed of: own land, savings, vehicle, fridge and TV. The housing quality index is composed of: pipped water, toilet, floor and roof materials.

[16] An alternative specification that clustered standard errors at the IFLS enumeration area yields qualitatively similar results (Table A.3).

Our empirical strategy relies on the identification assumption that exposure to pollution from the 1997 forest fires is uncorrelated with unobserved influences that vary across districts over time. Hence, if before the 1997 forest fires, districts that were more affected had differential growth in health outcomes compared to less affected districts, this could violate the identification assumption. To test for pre-trends before the fires, we use information on height in IFLS2 (surveyed in 1997) for children born between 1989 to 1994. Figure 2 shows the difference in height-for-age in affected and unaffected districts by year of birth and it is clear that there were no differences before the fires (Table A.4). Additionally, we assess the importance of selection and sorting on observable characteristics by examining the relationship between early-life exposure to the fires and individual and household observable characteristics. Table 2 presents the differences in the observed characteristics of exposed and non-exposed children. These differences are adjusted for district fixed effects and child's year and month of birth fixed effects. Exposure to the forest fires early in life is not significantly associated with the majority of socio-demographic characteristics except for land ownership.

4 Results

4.1 Short-term outcomes

We begin by analyzing children's short-term outcomes in 2000, 3 years after the fires, in Table 3. Early-life negative health shocks can affect children's physical development, which is captured here by height-for-age and stunting. We find that children who were exposed to the fires in their early life are shorter on average (columns 1-2). Among those exposed in-utero, they are on average 0.3 standard deviations shorter than unexposed children (columns 1-2). Children's exposure to the fires in their first and second year of life is also detrimental to their height, and the magnitudes are similar to in-utero exposure.[17] Regarding chronic malnutrition, children who were exposed in their second year of life are more likely to be stunted (columns 3-4), but children who were exposed in-utero or in their first year of life are

[17]We examine the effects by trimester and find that children exposed in their third trimester have lower height-for-age z-score, but the estimates are not significant for children exposed in their first or second trimester– this imprecise estimation is likely due to the small number of children exposed in each trimester of gestation. (Table C.5 Online Appendix)

not significantly more likely to be stunted in spite of their shorter stature.[18] These results suggest that children's health is affected in the short term, but it is not so severely affected that exposure to air pollution from the fires causes stunted growth.

4.2 Longer-term outcomes

Early-life negative health shocks disrupt children's human capital formation in the short-term, while there are two potential trajectories in the longer term. Children may be able to catch up, or investments and later interventions may mitigate the negative effects in the medium and long-terms. Alternatively, children may fail to catch up and continue to be disadvantaged later in life. Therefore, whether children's health continues to be affected due to exposure to acute pollution in early life poses an empirical question. We analyze children's medium-term outcomes 10 years post exposure (IFLS4 in 2007) in Table 4 and 17 years post exposure (IFLS5 in 2014) in Table 5. We begin with physical health status as measured by height and lung capacity, followed by cognitive function as measured by the Raven's test scores and mathematics test.

Ten years post exposure, we find that children who were exposed to the fires in-utero are on average 0.1 standard deviations (1.6 centimeters) shorter than children who were not exposed to the fires (Table 4, columns 1-2).[19] Children exposed to the fires in the first two years of life are also shorter, but the magnitude is smaller and not statistically significant.[20] These results suggest that children's physical development was affected 3 years after exposure in-utero, and the effect on height persists 10 years after exposure.

[18] We also explore the difference in height distribution and include some robustness checks. Figure A.2 (Panel A) shows the distribution of height-for-age z-score in 2000, 3 years post exposure. We find that the height distribution of children who were exposed to the fires in-utero is significantly different from the distribution of children who were not exposed to the fires. The joint test of significance for height-for-age z-score rejects the null that all coefficients are 0 (F-statistic 6.688, p-value=0.001). For robustness, we include district-specific linear trends and find a similar effect of in-utero exposure (-0.376, SE 0.181). We also conduct a permutation test and find that the p-value corresponding to the null hypothesis of no effect from exposure in-utero is 0.012. For exposure in the first and second year of life, the corresponding p-values are 0.016 and 0.010 respectively.

[19] We also explore differences in the height distribution of the exposed and non-exposed children 10 years post exposure. Figure A.2 (Panel B) shows the distribution of height in 2007. We find that the height distribution of children who were exposed to the fires in-utero is significantly different from the distribution of children who were not exposed to the fires at the 10 percent level. The joint test of significance for height rejects the null that all coefficients are 0 (F-statistic 2.745, p-value=0.051). For robustness, we include district-specific linear trends and find a similar effect of in-utero exposure to air pollution (-1.628 (SE 0.788)). We also conduct a permutation test and find that the p-value corresponding to the null hypothesis of no effect from exposure in-utero is 0.189.

[20] Using the CDC's height-for-age z-score for older children, we find that the height-for-age z-score of children who were exposed to the fires in-utero is lower than children who were not exposed to the fires. In 2007, the estimated effect is -0.225 (SE 0.097).

We expect exposure to air pollution to affect children's respiratory function contemporaneously; although, it is unclear whether the effect would persist 10 years after exposure. We find that children who were exposed to the fires in-utero and during their first 2 years of life have lower lung capacity than children who were not exposed to the fires (columns 3-4). The effects are between 5 to 6 percentage points, suggesting that the negative effect on respiratory function persists 10 years after the fires.

On the other hand, we find no statistically significant effect on cognitive function as measured by age-adjusted cumulative test score (cols. 5-6). Another supporting piece of evidence that corresponds to no effect on cognitive function is that children exposed to the fires are not delayed in starting school and they are on track for their age. These results imply that other factors might have mitigated the effects of the health shocks on cognitive function, or that the link between exposure to air pollution and cognitive function is not particularly strong in this case. More work on the link between air pollution and cognitive function would shed some light on the channels between air pollution and children's cognitive development.

We analyze the effects of early-life exposure to air pollution 17 years post-exposure in Table 5.[21] As adolescents, children exposed to the fires in-utero are likely to be 0.2 standard deviations (1.5 centimeters) shorter (columns 1-2).[22] These results are consistent with the persistent health effects of negative shocks in early life.[23] We find no statistically significant effect on respiratory function 17 years post-exposure (columns 3-4). Similar to our medium-term findings, we find no significant effect on children's cognitive function (columns 5-6) or the national examination test score in mathematics.[24]

The link between exposure to acute pollution and respiratory function does not persist

[21] We are underpowered to estimate the effect of acute air pollution in each trimester, but in general, the estimated effects of in-utero exposure to air pollution in the medium and longer-term are similar in each trimester (Table C.5 Online Appendix).

[22] We also explore differences in the height distribution of exposed and non-exposed children 17 years post exposure. Figure A.2 (Panel C) shows the distribution of height in 2014. While children who were exposed to the fires in-utero are on average shorter, the height distribution of children who were exposed to the fires in-utero is similar to the distribution of children who were not exposed to the fires. The joint test of significance for height fails to reject the null that all coefficients are 0 (F-statistic 1.558, p-value=0.209). For robustness, we conduct a permutation test and find that the p-value corresponding to the null hypothesis of no effect from exposure in-utero is 0.091. We also use height-for-age z-scores for older children and find that children who were exposed to the fires in-utero are on average -0.176 (SE 0.104) standard deviations shorter.

[23] One hypothesized link between in-utero exposure to air pollution and height is carbon monoxide that binds with hemoglobin, which in turn predisposes children to anemia. We explore this mechanism and find no statistically significant effect on hemoglobin count and anemia. The estimated effect on hemoglobin count is -0.020 (95% CI -0.38 to 0.34), while the estimated effect on anemia is 0.010 (95% CI -0.10 to 0.12). Another link between pollution and height is through early-life respiratory infections which we discuss in the next section.

[24] These results are consistent with a US-based study of children with asthma that finds no significant difference in test scores between children with and without asthma (Silverstein et al., 2001).

beyond 10 years post exposure. The change in the effect on respiratory function between 10 and 17 years post exposure may be due to children outgrowing respiratory symptoms. Our results are consistent with the literature that has found a gradual reduction in the incidence of wheezing and asthma-related symptoms as children transition into early adulthood (Bisgaard and Bønnelykke, 2010; Piippo-Savolainen and Korppi, 2008; Strachan, Butland and Anderson, 1996).

4.3 Potential channels

In this section, we explore the potential channels behind the persistent effects of exposure to pollution early-life. The negative shock to expectant mothers is one plausible channel underlying the effects found for children exposed in-utero. Exposure to the fires has been shown to affect adult health (Frankenberg, McKee and Thomas, 2005), so we specifically explore how exposure to the fires affected maternal health (Table 6, panel A, columns 1-2). The sample for this estimation includes ever married reproductive aged women between 15 and 49 years who were surveyed in the second wave of the IFLS in 1997 (IFLS2). IFLS2 allows us to analyze the immediate effects of the fires because more than 70% of the survey was conducted during the fires, between August and November 1997. We use geographical variation in exposure to the fires and differences in pregnancy status and estimate the following equation:

$$y_{id} = \delta pregnant_i \times affected_d + \gamma pregnant_i + \beta X_i + \psi_b + \rho_d + \epsilon_{id} \tag{2}$$

where y_i is the health status (in 1997) of woman i, residing in district d. The interactions $pregnant_i \times affected_d$ equals one if she is pregnant during the fires and resides in a district affected by the fires, while the indicator $pregnant_i$ takes the value one if she is pregnant at the time of the survey. X_i includes her education, urban residence, and (log) per capita expenditure. We include the women's year of birth fixed effects, ψ_b, and the 1997 district of residence fixed effects, ρ_d. Standard errors are clustered at the pollution grid level.

We analyze whether the mother's respiratory system, particularly her lung capacity, was affected by air pollution during pregnancy, which in turn may affect fetal development, including the child's respiratory system (Margulies, 1986; Murphy et al., 2005; Sbihi et al., 2016). We find that pregnant women exposed to the 1997 fires experience a 6 percentage point decrease in their lung capacity and a 7 percentage point increase in their self-reported breathing problems, which may explain the respiratory problems among children who were

exposed in-utero (Panel A, columns 1-2).[25] The fires may also affect women's nutritional status during pregnancy (Panel A, column 3). Specifically, we explore nutritional problems by analyzing the effects of the fires on expectant mother's body mass index (BMI). We find no significant difference among expectant mothers exposed and non-exposed to the fires, suggesting that women were able to obtain sufficient nutrition to gain weight during pregnancy in spite of the fires. These results suggest that the effect of the fires on maternal respiratory system during pregnancy is a likely channel behind children's subsequent health outcomes.

We then analyze the effects of the fires on children's health markers likely to respond immediately after the shock (Table 6, Panel B, columns 1 to 2). This estimation for the contemporaneous effects uses the second wave of the IFLS in 1997 (IFLS2). We include children who were born between 1995 and the time of the survey in 1997, so children in this sample are between 0 and 30 months at the time of the survey.[26] We exploit geographical variation in exposure to the fires and estimate the following equation:

$$y_{idmy} = \delta_1 exposed\_age1_{idmy} + \delta_2 exposed\_age2_{idmy} + \beta X_i + \tau_y + \mu_m + \rho_d + \epsilon_{idmy} \quad (3)$$

where y_{idmy} is the health status (in 1997) of child i, born in district d at month m and year y health status. The indicator $exposed\_age1_{idmy}$ ($exposed\_age2_{idmy}$) equals one if the child was exposed to the fires at any point in the first (second) year of life. The omitted category is children older than 24 months at the time of the survey. X_i includes maternal education, maternal age at delivery, urban residence, and (log) per capita expenditure. We include child birth year fixed effects, τ_y, and district of birth fixed effects, ρ_d. Standard errors are clustered at the pollution grid level.

We analyze the contemporaneous effects of air pollution on child health by focusing on the child's respiratory system and nutritional status. These are potential channels that link early-life exposure to air pollution and later outcomes. Our first potential channel is the incidence of respiratory problems within 4 weeks prior to the survey (IFLS2) as reported by the primary caregiver. During and immediately after the fires, we expect air pollution to affect children's respiratory system. Indeed, we find that children who were exposed

[25] If women of reproductive age who were exposed to the fires may experience a permanent health shock, then their children would not be a valid comparison group. First, we estimate pregnant women's lung capacity in affected districts after the fires in 2000 and find no statistically significant effects of pregnancy on respiratory function, the magnitude of the effect is 0.024 (0.019). Second, we find no statistically significant effect on women's lung capacity in 2000 among women who became pregnant immediately after the 1997 fires (0.026, SE 0.109). These results suggests that the effects of the fires do not seem to persist among women exposed to the fires in 1997 who became pregnant later.

[26] This sample is smaller than our sample in sub-sections 4.1 and 4.2 because some children who were exposed in-utero were not yet born by the time of the survey in 1997.

to the fires in their first two years of life are about 20 percentage points more likely to experience respiratory problems in 1997, during and immediately after exposure to the fires. This immediate effect on children's respiratory system appears to persist and is reflected in exposed children's lower lung capacity 10 years after the fires in 2007. The second potential channel is the contemporaneous child's nutritional status as measured by the probability of wasting, defined as less than 2 standard deviations in weight-for-height z-score. As shown in panel B, we find no strong evidence in line with this channel. Taken together, this pattern of strong contemporaneous effects on children's respiratory function and insignificant evidence on immediate nutritional status is consistent with the hypothesis that the persistent effect on height is driven by the link between early-life respiratory infections and later physical growth.

Given the potential link between maternal health status and in-utero growth, we also explore whether in-utero exposure to the fires affected birth weight (Panel B, column 3). Birth weight is often used in the literature as a proxy for health endowments at birth, birth weight may reflect prenatal inputs and allow us to capture the effects of exposure to air pollution in-utero (Bharadwaj, Eberhard and Neilson, 2010; Rosenzweig and Zhang, 2009).[27] We include children who were born between 1995 and 2000 and estimate the following equation:

$$y_{idmy} = \delta_0 inutero_{idmy} + \beta X_i + \tau_y + \mu_m + \rho_d + \epsilon_{idmy} \tag{4}$$

where y_{idmy} is birth weight in kilograms (based on mother's recall), *inutero* equals one if the child was exposed to the fires in-utero, X_i includes maternal education, maternal age, urban residence, and (log) per capita expenditure. We include child birth and month year fixed effects, τ_y, μ_m and district of birth fixedh effects, ρ_d. Standard errors are clustered at the pollution grid level. In-utero exposure to the fires has no statistically significant effect on birth weight. The observed long-term effects appear to be driven by maternal respiratory health and children's early life respiratory health.

We also analyze how some health investments may have responded to exposure to early-life pollution. In particular, we examine breastfeeding, as measured by the probability of breastfeeding for at least 6 months, and fertility decisions, as measured by spacing between the index child and the subsequent child (Panel C, columns 1 and 2 respectively). We estimate equation 1 with health investments as dependent variables for the children in our

[27] Low birth weight (birth weight $<$ 2,500 grams) is an important health outcome because it is correlated with later outcomes, but there may be measurement error in birth weight reporting. Birth weight is available for only about 70% of all children and there is evidence of 'lumping' in 500-gram increments.

main sample (those born between 1995 and 2000).[28] On breastfeeding, it is possible that mothers who were less healthy due to the fires would have to stop breastfeeding earlier than healthy mothers. On the other hand, it is probable that mothers might increase breastfeeding to improve their children's immune system. We find no statistically significant effect on breastfeeding behavior associated with a child's exposure to the fires. Additionally, birth spacing may increase if women's health were affected by the fire, or they may decide to wait longer before conceiving another child to invest more on the child experiencing the negative health shock. We find that affected children are not significantly more likely to have younger siblings sooner than non-affected children.

4.4 Threats to Identification and Robustness

Falsification tests Our main threat to identification is that exposure to pollution may be spurious and confounded with unobserved shocks or trends. In this section, we provide some evidence that validate our empirical strategy. To begin, negative shocks several months before conception should not have a causal impact on birth outcomes. Table 7 presents 'falsification regressions' that include exposure to air pollution six months before conception as a way to check if there were spurious differential trends between children in places more and less affected by the pollution shock. If our estimated effects of early-life exposure to the fires on children's human capital outcomes in the short and long-term were biased by the presence of omitted variables or a trend, then these placebo regressions may show significant effects as well. Our results show no impact of these prior exposures and no change in the magnitude of the effects of exposure in-utero and at ages 1 and 2 years.

We also perform several analyses to alternatively test the parallel trends assumption underlying our difference-in-difference estimation. Our estimated model depicted in equation 1 implies that the younger cohorts are part of the comparison group. Therefore, in the absence of the 1997 fires, the average difference in outcomes between children born in 1995 to early 1998 in affected and unaffected districts should be the same as the difference for younger cohorts (born in late 1998 to 2000). Appendix Table A.5 estimates a version of equation 1 that interacts a dummy of affected districts with the cohorts in our sample. These results show that there is no effect for the non-treated cohorts (i.e the cohort conceived one year after the fires (born in 1999), where the cohort conceived two years after (born in

[28] Children who were exposed to the fires at 13 and 24 months old may no longer be breastfed. In order to take this into account, we also estimated the following equation: $y_{idmy} = \delta_0 inutero_{idmy} + \delta_1 age1_{idmy} + \beta X_i + \tau_y + \mu_m + \rho_d + \epsilon_{idmy}$where y_{idmy} is being breast-fed for more than 6 months (based on mother's recall), $inutero\,(age1)$ equals one if the child was exposed to the fires in-utero (age 1), X includes maternal education, maternal age, urban residence, and (log) per capita expenditure. Exposure to the fires has no statistically significant effect on the probability of being breastfed on this alternative sample.

2000) is the omitted group).[29] In addition, we perform a falsification test that also serve to test the parallel trend assumption: we move the fires dates to later years to create a "fake fires" dates. We moved the fires dates to August to November 2003 to check if there is any difference in human capital outcomes between children in affected and unaffected districts after exposure to the fake fires, which could detect any spurious differential trends. Appendix Table A.6 shows the results of estimating equation 1 for a sample of children born between 2001 and 2006 using human capital outcomes in IFLS5 in 2014. We use this sample because these children are similar in age to children in our main analyzed sample in 2007, who were between 7 to 12 in IFLS4. We find no evidence of effects of exposure to the fake fires.

Robustness We explore the role of other household shocks as potential confounders. We analyze the role of the local economy by taking into account the 1997 economic crisis and crop loss due to the El Niño drought (table A.7). We include the change in the province unemployment in the child's year of birth, self-reported crop loss and self-reported income loss in the 1997 survey to proxy for the effects of the economic crisis. We find that children who were exposed to the fires in-utero have poorer health outcomes as measured by height in the short and longer-term, and the estimated effect sizes are similar to our earlier results, thereby suggesting that the effects cannot be attributed to the Asian economic crisis.

As specification tests, we show that our results are robust to alternative definitions of the shock, and explore different thresholds of the aerosol index for sensitivity. First, Table C.1 (Online appendix) shows that our results are robust to redefining the exposure threshold using a slightly higher (aerosol greater that 1.75) or lower (aerosol greater that 1.25) cutoff. Second, we explore non-linearities in two dimensions. 1) We use aerosol index above 1.5 as the cutoff and vary the threshold of exposure at 0 days (reference category), 1-2 days, 3-6 days, and 7 days or more (Table C.3 in the Online appendix). 2) We use 3 days or more of exposure as the cutoff and vary the aerosol index threshold at 0 to 1 (reference category), 1 to 1.5, 1.5 to 2.25 and above 2.25 (Table C.4 in the Online appendix). These alternative analyses provide evidence of the non-linear effects of exposure to air pollution. Worse later life outcomes appear to be driven by exposure to higher pollution days and higher pollution levels.[30]

Another concern is the possibility of measurement error in our treatment variable. Ex-

[29] We estimate the following equation: $y_{idmy} = \Sigma_{y=age\_2}^{y=age\_-2} Affected\_Dis_d * cohort_y + \beta X_i + \tau_y + \mu_m + \rho_d + \epsilon_{idmy}$, where our coefficients of interest are the interactions between being born in a district affected by the 1997 forest fires ($Affected\_Dis_d$) and cohort: 25-36 months during the fires, 13-24 months during the fires, 0-12 months during the fires, in-utero during the fires, or conceived 1 year after. Figure A.3 presents the estimates for height in 2014.

[30] We also explore a continuous measure, defined as the mean aerosol index in-utero, during 0-12 months and 13-24 months in Online Appendix Table C.2.

posure to pollution from the forest fires is measured by the aerosol index, which may result in noisy measures for households far from the TOMS pollution centroid grids compared to households closer to the centroids. Since we lack information on household location, we rely on information on the district of birth to determine exposure to the 1997 forest fires. To explore the potential problem of measurement error, we restrict the distance between the district of birth centroid and TOMS grids to the 50th percentile, corresponding to 32 kilometers, and the 75th percentile, corresponding to 38 kilometers. The results are robust and similar in magnitude to the full sample estimates, which suggests that the potential of mis-measurement of pollution should not pose a serious concern in the main analysis (Figure A.4).

Selection issues One may be concerned that exposure to the forest fires affected women's fertility and migration decisions, which could confound our estimates of early-life exposure to pollution on children's later-life outcomes. In this section, we address these concerns by analyzing migration and fertility responses to exposure to the 1997 fires. Regarding fertility, it is ambiguous whether more advantaged or disadvantaged women conceive during the forest fires or delay pregnancy, which could weaken the validity of using younger cohorts as part of the comparison group. For example, if more disadvantaged women conceived during the fires, part of the treatment group could be negatively selected. In order to assess whether some women are more or less likely to conceive during or after the fires, we estimated models of the following form using a sample of women in fertile ages:

$$M_{idp} = \delta affected_d + \gamma affected_d * X_i + \beta X_i + \rho_p + \varepsilon_{idp} \tag{5}$$

where M_{idp} measures two fertility indicators: 1) if woman i in district d and province p was pregnant during the 1997 forest fires months, and 2) if she conceived in the following 12 months after the forest fires. $affected_d$ is an indicator of whether district d was affected by the 1997 forest fires and we interact this indicator with women's socio-demographic characteristics to assess the degree of differential responses by observable characteristics. We include province fixed effects instead of district fixed effects because this regression relies exclusively on geographic variation.[31] Appendix table A.8 shows that there is little evidence of endogenous fertility responses both during (Column 1) or after the forest fires (Column 2).

Next, one may be concerned that exposure to the fires influenced families' migration decisions and that such migration responses may have changed the socio-demographic com-

[31] These regression models lack the cohort variation we exploited in the children's regressions because we rely on pregnancy status during 1997.

position of families living in places more or less affected by the fires, thereby confounding the true effect of exposure to air pollution. In column 3 of Table A.8, we examine whether the probability of migrating after 1997 was different in places affected by the fires and the degree of heterogeneous responses by observable characteristics using a similar specification as equation 5. We do not find evidence that migration responses were different between districts exposed and not exposed to the fires.

5 Discussion and Conclusion

This paper examines the short and long-term effects of early-life exposure to air pollution in a low middle-income country. We find persistent effect on height and lung capacity among children who were exposed in-utero to the 1997 Indonesian fires. Additionally, these persistent effects may be due to expectant mothers and young children experiencing poorer respiratory health during the fires. These results are consistent with the hypothesis that the long-term effect on height is driven by the link between early-life respiratory infections and later physical growth.

Exposure to air pollution during the fires may affect children with different backgrounds differently, so we explore heterogeneous treatment effects. We begin with gender, since the fragile male hypothesis argues that males are more vulnerable to shocks in early life compared to females (Kraemer, 2000). According to this hypothesis, males would be more affected by early-life exposure to air pollution. We find no evidence of heterogeneity by gender in the short and longer-term among children who were exposed in-utero (Online Appendix Table B.1). More importantly, we find no evidence of heterogeneity by socio-economic characteristics: maternal education, urban residence, poverty (defined as the bottom two quintiles of per capita expenditure), agricultural participation, and chronic indoor pollution from wood-burning stove (Online Appendix Tables B.2 to B.6). The adverse effects of acute exposure to air pollution on children's health are indiscriminate, which highlights the difficulty in escaping the effects of a severe negative shock like these fires.

Exposure to the Indonesian fires can be compared to the effects of other forms of outdoor pollution and even chronic indoor pollution such as burning biomass fuels. During the 1997 Indonesian fires, levels of particulate matter (TSP) ranged between 150 and 4000 $\mu g/m^3$, with PM_{10} levels surpassing 2000 $\mu g/m^3$ at the peak. For chronic indoor pollution, estimated levels of PM_{10} associated with biomass fuels range between 1000 and 2000 $\mu g/m^3$ at peak indoor concentration (Fullerton, Bruce and Gordon, 2008). The estimated effects of exposure to burning organic matter from the forest fires are based on approximately 15 days of exposure and these effects persist in children's respiratory function in childhood and

shorter stature in their teenage years. Because of the similar source of pollution from burning organic matter through biomass fuels or forest fires, the evidence from our study can inform the long-term effects of chronic indoor pollution.

Our estimates of the long-term effects of early-life exposure to pollution come from surviving children, so observed outcomes combine the effects of culling and scarring. Jayachandran (2009) finds that the fires were associated with a 1.2 percent reduction in birth cohort, suggesting fetal death and infant mortality. In the presence of early life mortality, we do not observe the long-term effects of the children who died due to culling. If the selection effect from culling dominates the scarring effect, surviving children may not exhibit adverse long term outcomes. On the other hand, if the scarring effect dominates, we observe adverse outcomes among survivors. Our results suggest persistent effects on lung capacity and height among survivors, which is consistent with the scarring effect dominating culling. Therefore, our results could be interpreted as a lower bound.[32]

The persistent effects on children's height may be associated with later earnings (Case and Paxson, 2008; Persico, Postlewaite and Silverman, 2004; Schultz, 2002; Thomas and Strauss, 1997; Vogl, 2014). If we use the estimated height premium of 8 percent for an additional 10 centimeters in physical stature for Indonesia (Sohn, 2015), and that this premium is not mediated by differences in cognitive function (LaFave and Thomas, 2013), the estimated effect of almost 2 centimeters lower height from exposure to the fires would translate to 1.6 percent lower earnings for the affected population. Based on our sample, about 10 percent of surviving children under 2 years were affected by the fires. Therefore, we assume that this fraction of the workforce would be about 2 centimeters shorter, and earn less once they join the labor market. Extrapolating these estimates, at Indonesia's current GDP of USD 868.2 billion and assuming that 10 percent of the workforce were affected by the fires, the effect of lower height on earnings translates to 1.4 billion lower GDP (0.15 percent of GDP) annually. The initial cost of the fires is about 2 percent of GDP, and the long-term effect is 7.5 percent of the initial loss annually. These estimates suggest that the hidden costs of air pollution can be long-lasting and substantial.

The annual occurrence of forest fires in Indonesia continues to be a policy challenge for Southeast Asia. In fact, the most recent 2015 fires were similarly exacerbated by the El Niño weather phenomenon, like the 1997 episode. Due to climate change, these events are more likely to occur, not only in developing countries but also in high income countries. Therefore, interventions that reduce the occurrence of uncontrolled fires are necessary.

[32] Jayachandran (2009) finds a 1.2 percent decrease in cohort size due to prenatal exposure to the fires. Using back of the envelope calculations, if we assumed that the children who died had survived, and that their height 17 years post-exposure would have been in the bottom 1.2 percentile of survivors' height, our estimated effect would increase from 1.55 to 1.86 centimeters.

Additionally, complementary policy options would involve developing interventions that encourage avoidance behavior. For instance, the government may consider developing an air quality warning system to limit exposure to air pollution, especially for young children and expectant mothers.

References

Alderman, Harold, John Hoddinott, and Bill Kinsey. 2006. "Long term consequences of early childhood malnutrition." *Oxford economic papers*, 58(3): 450–474.

Almond, Douglas, and Bhashkar Mazumder. 2013. "Fetal origins and parental responses." *Annual review of economics*, 5: 37–56.

Almond, Douglas, and Janet Currie. 2011. "Killing me softly: The fetal origins hypothesis." *The Journal of Economic Perspectives*, 25(3): 153–172.

Almond, Douglas, Lena Edlund, and Mårten Palme. 2009. "Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden." *The Quarterly journal of economics*, 124(4): 1729–1772.

Almond, Douglas, Yuyu Chen, Michael Greenstone, and Hongbin Li. 2009. "Winter Heating or Clean Air? Unintended Impacts of China's Huai River Policy." *American Economic Review*, 99(2): 184–90.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City." *The Economic Journal*.

Bailey, Roy E, Timothy J Hatton, and Kris Inwood. 2016. "Atmospheric Pollution and Child Health in Late Nineteenth Century Britain."

Banerjee, Rakesh. 2016. "Long Run Effects of Exposure to Forest Fires in Indonesia."

Barker, David J. 1995. "Fetal origins of coronary heart disease." *BMJ: British Medical Journal*, 311(6998): 171.

Bharadwaj, Prashant, Juan Eberhard, and Christopher Neilson. 2010. "Do initial endowments matter only initially? The persistent effect of birth weight on school achievement."

Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher Neilson. 2017. "Gray matters: fetal pollution exposure and human capital formation." *Journal of the Association of Environmental and Resource Economists*, 4(2): 505–542.

Bisgaard, Hans, and Klaus Bønnelykke. 2010. "Long-term studies of the natural history of asthma in childhood." *Journal of Allergy and Clinical Immunology*, 126(2): 187–197.

Black, Sandra, Aline Bütikofer, Paul J Devereux, and Kjell G Salvanes. 2013. "DP9443 This Is Only a Test? Long-Run Impacts of Prenatal Exposure to Radioactive Fallout."

Bobak, Martin, Marcus Richards, and Michael Wadsworth. 2004. "Relation between children's height and outdoor air pollution from coal-burning sources in the British 1946 birth cohort." *International archives of occupational and environmental health*, 77(6): 383–386.

Case, Anne, and Christina Paxson. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy*, 116(3).

Chay, Kenneth Y, and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics*, 118(3).

Cunha, Flavio, and James Heckman. 2007. "The economics of human development: The technology of skill formation." *The American economic review*, 97(2): 31–47.

Currie, Janet, and Tom Vogl. 2013. "Early-Life Health and Adult Circumstance in Developing Countries." *Annu. Rev. Econ*, 5: 1–36.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know About Short-and Long-Term Effects of Early-Life Exposure to Pollution?" *Annu. Rev. Resour. Econ*, 6: 1–1.

Currie, Janet, Matthew Neidell, and Johannes F Schmieder. 2009. "Air pollution and infant health: Lessons from New Jersey." *Journal of health economics*, 28(3): 688–703.

Dixon, John A, David P Gibbon, and Aidan Gulliver. 2001. *Farming systems and poverty: improving farmers' livelihoods in a changing world.* Food & Agriculture Org.

Emmanuel, Shanta Christina. 2000. "Impact to lung health of haze from forest fires: the Singapore experience." *Respirology*, 5(2): 175–182.

Fink, Günther, Isabel Günther, and Kenneth Hill. 2014. "Slum residence and child health in developing countries." *Demography*, 51(4): 1175–1197.

Foster, Andrew, Emilio Gutierrez, and Naresh Kumar. 2009. "Voluntary compliance, pollution levels, and infant mortality in Mexico." *The American Economic Review*, 99(2): 191–197.

Frankenberg, Elizabeth, Douglas McKee, and Duncan Thomas. 2005. "Health consequences of forest fires in Indonesia." *Demography*, 42(1): 109–129.

Fullerton, Duncan G, Nigel Bruce, and Stephen B Gordon. 2008. "Indoor air pollution from biomass fuel smoke is a major health concern in the developing world." *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 102(9): 843–851.

Ghosh, Rakesh, E Amirian, Miroslav Dostal, Radim J Sram, and Irva Hertz-Picciotto. 2011. "Indoor coal use and early childhood growth." *Archives of pediatrics & adolescent medicine*, 165(6): 492–497.

Giglio, Louis, James T Randerson, and Guido R Werf. 2013. "Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4)." *Journal of Geophysical Research: Biogeosciences*, 118(1): 317–328.

Greingor, JL, JM Tosi, S Ruhlmann, and M Aussedat. 2001. "Acute carbon monoxide intoxication during pregnancy. One case report and review of the literature." *Emergency Medicine Journal*, 18(5): 399–401.

Hanna, Rema, and Paulina Oliva. 2016. "Implications of Climate Change for Children in Developing Countries." *The Future of Children*, 115–132.

Hanna, Rema, Esther Duflo, and Michael Greenstone. 2016. "Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves." *American Economic Journal: Economic Policy*, 8(1): 80–114.

Heil, Angelika. 1998. "Air Pollution Caused by Large Scale Forest Fires in Indonesia 1997."

Heil, Angelika, and Johan Goldammer. 2001. "Smoke-haze pollution: a review of the 1997 episode in Southeast Asia." *Regional Environmental Change*, 2(1): 24–37.

Heinrich, Joachim, and Remy Slama. 2007. "Fine particles, a major threat to children." *International journal of hygiene and environmental health*, 210(5): 617–622.

Isen, Adam, Maya Rossin-Slater, and W Reed Walker. 2014. "Every Breath You Take-Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970." National Bureau of Economic Research.

Jayachandran, Seema. 2009. "Air quality and early-life mortality evidence from Indonesia's wildfires." *Journal of Human Resources*, 44(4): 916–954.

Kajekar, Radhika. 2007. "Environmental factors and developmental outcomes in the lung." *Pharmacology & therapeutics*, 114(2): 129–145.

Koplitz, Shannon N, Loretta J Mickley, Miriam E Marlier, Jonathan J Buonocore, Patrick S Kim, Tianjia Liu, Melissa P Sulprizio, Ruth S DeFries, Daniel J Jacob, Joel Schwartz, et al. 2016. "Public health impacts of the severe haze in Equatorial Asia in September–October 2015: demonstration of a new framework for informing fire management strategies to reduce downwind smoke exposure." *Environmental Research Letters*, 11(9): 094023.

Kraemer, Sebastian. 2000. "The fragile male." *British Medical Journal*, 321(7276): 1609.

Kyu, Hmwe Hmwe, Katholiki Georgiades, and Michael H Boyle. 2009. "Maternal smoking, biofuel smoke exposure and child height-for-age in seven developing countries." *International journal of epidemiology*, 38(5): 1342–1350.

LaFave, Daniel, and Duncan Thomas. 2013. "Height and cognition at work: labor market performance in a low income setting." *Duke University*.

Landrigan, Philip J, Richard Fuller, Nereus JR Acosta, Olusoji Adeyi, Robert Arnold, Abdoulaye Bibi Baldé, Roberto Bertollini, Stephan Bose-O'Reilly, Jo Ivey Boufford, Patrick N Breysse, et al. 2017. "The Lancet Commission on pollution and health." *The Lancet*.

Magnusson, Patrik KE, Finn Rasmussen, and Ulf B Gyllensten. 2006. "Height at age 18 years is a strong predictor of attained education later in life: cohort study of over 950 000 Swedish men." *International Journal of Epidemiology*, 35(3): 658–663.

Margulies, Jeffrey L. 1986. "Acute carbon monoxide poisoning during pregnancy." *The American journal of emergency medicine*, 4(6): 516–519.

McPhaden, MJ. 2015. "Playing hide and seek with El Niño." *Nature Climate Change*, 5(9): 791–795.

Mishra, Vinod, and Robert D Retherford. 2007. "Does biofuel smoke contribute to anaemia and stunting in early childhood?" *International Journal of Epidemiology*, 36(1): 117–129.

Mortimer, Kathleen, Romain Neugebauer, Frederick Lurmann, Siana Alcorn, John Balmes, and Ira Tager. 2008. "Air pollution and pulmonary function in asthmatic children: effects of prenatal and lifetime exposures." *Epidemiology*, 19(4): 550–557.

Murphy, VE, PG Gibson, R Smith, and VL Clifton. 2005. "Asthma during pregnancy: mechanisms and treatment implications." *European Respiratory Journal*, 25(4): 731–750.

Neidell, Matthew J. 2004. "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *Journal of health economics*, 23(6): 1209–1236.

Ostermann, Kathryn, and Michael Brauer. 2001. "Air quality during haze episodes and its impact on health." *Forest Fires and Regional Haze in Southeast Asia, ed. by P. Eaton, and M. Radojevic.*

Peet, Evan D. 2015. "Environment and Human Capital: The Effects of Early-Life Exposure to Pollutants in the Philippines."

Persico, Nicola, Andrew Postlewaite, and Dan Silverman. 2004. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *Journal of Political Economy*, 112(5).

Piippo-Savolainen, Eija, and Matti Korppi. 2008. "Wheezy babies–wheezy adults? Review on long-term outcome until adulthood after early childhood wheezing." *Acta paediatrica*, 97(1): 5–11.

Plopper, Charles G, and Michelle V Fanucchi. 2000. "Do urban environmental pollutants exacerbate childhood lung diseases?" *Environmental health perspectives*, 108(6): A252.

Rangel, Marcos A, and Tom Vogl. 2016. "Agricultural Fires and Infant Health."

Rona, Roberto J. 1981. "Genetic and environmental factors in the control of growth in childhood." *British Medical Bulletin*, 37(3): 265–272.

Rosenzweig, Mark R, and Junsen Zhang. 2009. "Do population control policies induce more human capital investment? Twins, birth weight and China's one-child policy." *The Review of Economic Studies*, 76(3): 1149–1174.

Rückerl, Regina, Alexandra Schneider, Susanne Breitner, Josef Cyrys, and Annette Peters. 2011. "Health effects of particulate air pollution: a review of epidemiological evidence." *Inhalation toxicology*, 23(10): 555–592.

Sanders, Nicholas J. 2012. "What doesn't kill you makes you weaker prenatal pollution exposure and educational outcomes." *Journal of Human Resources*, 47(3): 826–850.

Sastry, Narayan. 2002. "Forest fires, air pollution, and mortality in Southeast Asia." *Demography*, 39(1): 1–23.

Sbihi, Hind, Lillian Tamburic, Mieke Koehoorn, and Michael Brauer. 2016. "Perinatal air pollution exposure and development of asthma from birth to age 10 years." *European Respiratory Journal*, ERJ–00746.

Schultz, T Paul. 2002. "Wage Gains Associated with Height as a Form of Health Human Capital." *American Economic Review*, 92(2): 349–353.

Sheng, Tan Jie. 2017. "Health Irreversibility from Early-Life Exposure to Air Pollution: Evidence from Four Waves of Indonesian Longitudinal Data."

Silverstein, Marc D, Joanne E Mair, Slavica K Katusic, Peter C Wollan, Edward J O'Connell, and John W Yunginger. 2001. "School attendance and school performance: a population-based study of children with asthma." *The Journal of pediatrics*, 139(2): 278–283.

Silwal, Ani Rudra, and Andy McKay. 2014. "The Impact of Cooking with Firewood on Respiratory Health: Evidence from Indonesia." Department of Economics, University of Sussex.

Sohn, Kitae. 2015. "The height premium in Indonesia." *Economics & Human Biology*, 16: 1–15.

Strachan, David P, Barbara K Butland, and H Ross Anderson. 1996. "Incidence and prognosis of asthma and wheezing illness from early childhood to age 33 in a national British cohort." *Bmj*, 312(7040): 1195–1199.

Tacconi, Luca. 2003. "Fires in Indonesia: causes, costs and policy implications." CIFOR, Bogor, Indonesia.

Thomas, Duncan, and John Strauss. 1997. "Health and wages: Evidence on men and women in urban Brazil." *Journal of Econometrics*, 77(1): 159–185.

Torres, O, PK Bhartia, JR Herman, A Sinyuk, Paul Ginoux, and Brent Holben. 2002. "A long-term record of aerosol optical depth from TOMS observations and comparison to AERONET measurements." *Journal of the atmospheric sciences*, 59(3): 398–413.

Ungvary, G, and E Tatrai. 1985. "On the embryotoxic effects of benzene and its alkyl derivatives in mice, rats and rabbits." In *Receptors and Other Targets for Toxic Substances*. 425–430. Springer.

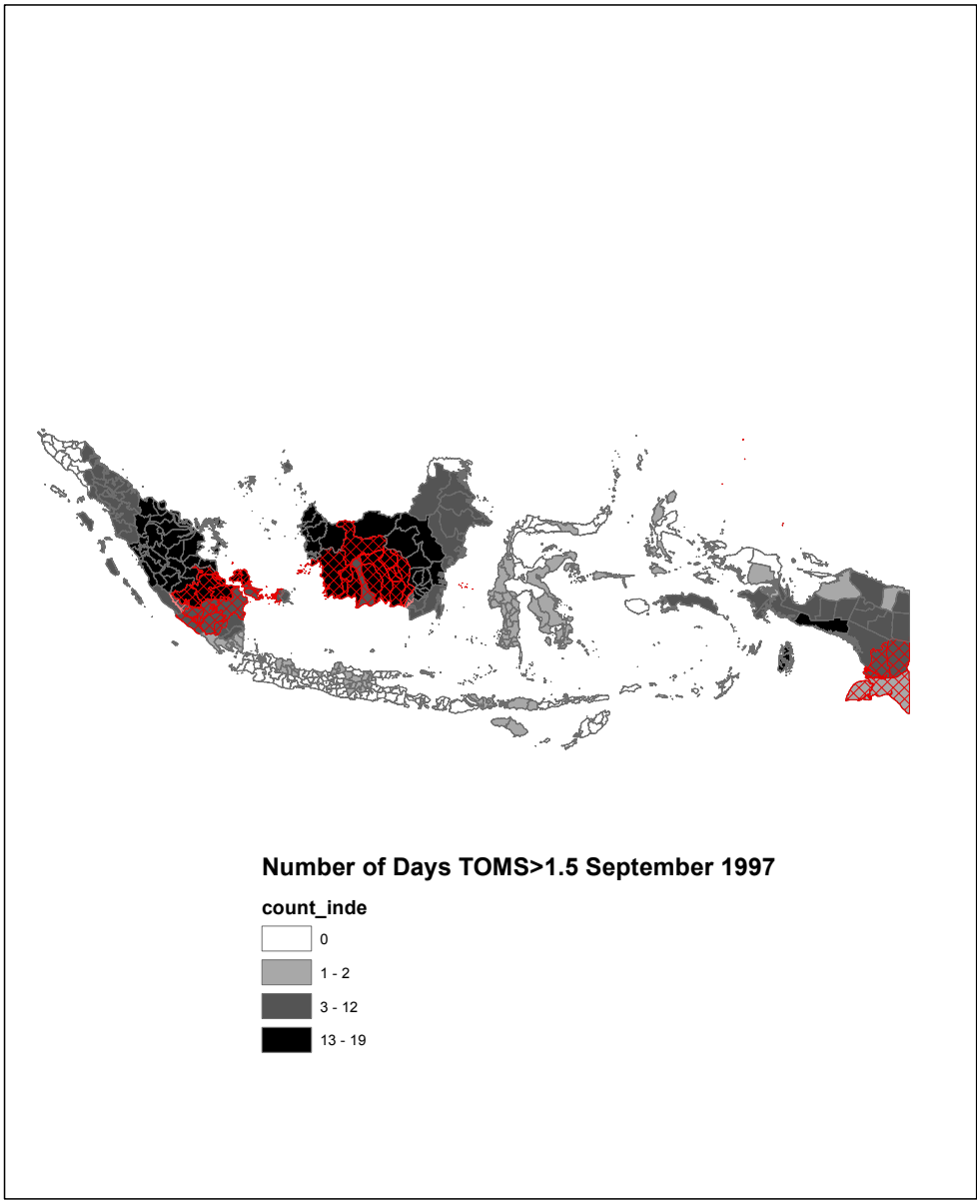
United Nations Framework Convention on Climate Change. 2007. "Climate Change: Impacts, Vulnerabilities, and Adaptation in Developing Countries." United Nations Framework Convention on Climate Change Bonn.

Vik, Torstein, Geir Jacobsen, Lars Vatten, and Leiv S Bakketeig. 1996. "Pre- and post-natal growth in children of women who smoked in pregnancy." *Early human development*, 45(3): 245–255.

Vogl, Tom S. 2014. "Height, skills, and labor market outcomes in Mexico." *Journal of Development Economics*, 107: 84–96.

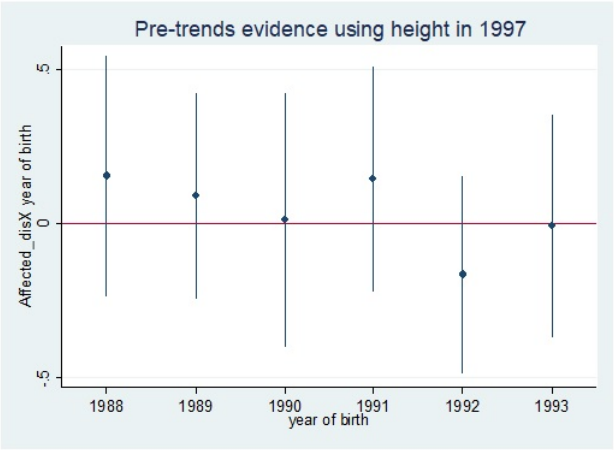
Tables and Figures

Figure 1: Burnt Area and TOMS Aerosol



Notes: Aerosol index comes from TOMS dataset and burnt area (in hectares) comes from the fourth version of the Global Fire Emissions Database (GFED4). Pollution exposure is defined the number of days that the aerosol index exceeds 1.5. Burnt area is depicted in red.

Figure 2: Pre-trends based on older children's height



Note: Coefficients on the interaction between year of birth and affected areas. All standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table 1: Summary statistics

	(1)	(2)	(3)	(4)
	Non-exposed		Exposed	
	Mean	N	Mean	N
Panel A. Short term outcomes: 3 years post fires				
Height for age	-1.464	3,572	-2.033	413
	(2.291)		(1.404)	
Stunted	0.378	3,509	0.464	407
	(0.485)		(0.499)	
Panel B. 10 years post fires				
Height (cm)	127.88	3,833	133.17	447
	(11.89)		(9.33)	
Lung capacity (log)	5.339	2,286	5.332	447
	(0.290)		(0.254)	
Total test score	0.695	3,052	0.713	406
(fraction correct)	(0.177)		(0.174)	
Panel C. 17 years post fires				
Height (cm)	156.88	3,360	157.33	368
	(8.39)		(8.42)	
Lung capacity (log)	5.840	3,345	5.916	366
	(0.273)		(0.267)	
Total test score	0.737	2,807	0.745	293
(fraction correct)	(0.175)		(0.166)	

Notes: The sample is restricted to children born between 1995 and 2000. Exposed children include children who were exposed to an aerosol index of more than 1.5 for at least 3 days in a month when the child is in-utero, between 0-12 months, or 13 to 24 months. Non-exposed children include all other children born in areas affected by the fires or children who were born in areas not affected by the fires. The aerosol index ranges from 0 to almost 6 in our sample. Children's respiratory problem in the past four weeks prior to the survey is reported by the mother or primary caretaker. Stunted takes the value 1 if a child is less than 2 standard deviations below average for height-for-age. Lung capacity is the maximum of the three readings taken in the survey. This measure is only available for respondents over the age of 9. Total test score is given by the fraction correct on the Raven progressive matrices and mathematics questions. IFLS2 was conducted during or immediately after the fires in 1997, IFLS3 was conducted 3 years after the fires in 2000, IFLS4 was conducted 10 years after the fires in 2007, IFLS5 was conducted 17 years after the fires in 2014.

Table 2: Balance test: differences in the characteristics of exposed and non-exposed children

	(1) Male child	(2) Household size	(3) Poor household	(4) Own land	(5) Urban	(6) Mother's education: primary	(7) Mother's BMI	(8) Mother's lung capacity (log)
In-utero	-0.051	0.109	0.067	0.029	0.005	0.053	-0.045	-9.897
	(0.063)	(0.265)	(0.057)	(0.050)	(0.028)	(0.038)	(0.362)	(7.197)
0-12mo	0.014	-0.017	0.017	-0.116**	-0.024	-0.062	0.296	-3.279
	(0.053)	(0.256)	(0.048)	(0.049)	(0.042)	(0.048)	(0.392)	(6.030)
13-24mo	-0.031	0.017	-0.024	0.024	-0.036	0.013	-0.165	-4.340
	(0.050)	(0.227)	(0.053)	(0.062)	(0.034)	(0.036)	(0.389)	(5.921)
Obs.	3,790	3,790	3,790	3,790	3,790	3,790	3,672	3,661
R-sq.	0.053	0.112	0.135	0.126	0.564	0.225	0.074	0.201

Notes: Mother's primary education takes the value one if she has primary education or lower. Poor takes the value one if the household is in the bottom two quintiles of total per capita expenditure in 1997 (1 USD˜Rp. 10,000). The regressions include district, children's year and month of birth fixed effects. All standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Short-term outcomes: 3 years post-exposure

	(1)	(2)	(3)	(4)
	Height for age z-score		Stunted	
In-utero	-0.315	-0.354**	-0.038	-0.036
	(0.200)	(0.177)	(0.054)	(0.051)
0-12mo	-0.387***	-0.386***	0.061	0.057
	(0.107)	(0.104)	(0.053)	(0.051)
13-24mo	-0.252***	-0.253***	0.081**	0.080**
	(0.088)	(0.089)	(0.039)	(0.039)
Observations	3,790	3,790	3,790	3,790
R-squared	0.187	0.219	0.152	0.175
HH char.	N	Y	N	Y
Mean of	-1.464		0.378	
non-exposed	(2.291)		(0.485)	

Notes: Height-for-age z-score measured in wave 3 (2000). Stunted takes the value one if child's height-for-age z-score is more than two standard deviations below the mean. Covariates include male child, mother's age at delivery, and mother's education, urban residence, household size, household per capita expenditure (log), an asset index and a housing quality index. Mother's primary education takes the value one if she has primary education or lower. Per capita expenditure is in 2000 Rupiah (1 USD˜Rp. 10,000 in 2000). The regressions include district, children's year and month of birth fixed effects. All standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Children's outcomes in 2007: 10-years post-exposure

	(1)	(2)	(3)	(4)	(5)	(6)
	Height		Lung capacity		Cognitive test	
In-utero	-1.819***	-1.619**	-0.062**	-0.057*	-0.062	-0.053
	(0.663)	(0.622)	(0.030)	(0.033)	(0.135)	(0.127)
0-12mo	-0.902	-0.657	-0.052**	-0.049**	-0.015	-0.006
	(0.757)	(0.723)	(0.022)	(0.023)	(0.078)	(0.080)
13-24mo	-0.366	-0.156	-0.060***	-0.053***	0.115	0.114
	(0.729)	(0.708)	(0.019)	(0.020)	(0.105)	(0.097)
Observations	4,180	4,180	2,386	2,386	3,344	3,344
R-squared	0.669	0.682	0.361	0.381	0.137	0.192
HH char.	N	Y	N	Y	N	Y
Mean of non-exposed	127.88		5.339		0.008	
group	(11.89)		(0.290)		(0.996)	

Notes: The maximum of 3 measures of lung capacity is used (in log). Age-adjusted test score is the age-adjusted z-score of the fraction correct response on both the Raven and mathematics tests. Covariates include male child, mother's age at delivery, and mother's education, urban residence, household size, household per capita expenditure (log), an asset index and a housing quality index. Mother's primary education takes the value one if she has primary education or lower. Per capita expenditure is in 2000 Rupiah (1 USD˜Rp. 10,000 in 2000. The regressions include district, children's year and month of birth fixed effects. All standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Children's outcomes in 2014: 17 years post-exposure

	(1)	(2)	(3)	(4)	(5)	(6)
	Height		Lung capacity		Cognitive test	
In-utero	-1.451*	-1.556**	0.036	0.035	0.076	0.103
	(0.799)	(0.774)	(0.026)	(0.026)	(0.126)	(0.124)
0-12mo	-0.527	-0.356	0.022	0.025	-0.027	-0.011
	(0.752)	(0.743)	(0.019)	(0.019)	(0.088)	(0.082)
13-24mo	-0.893	-0.741	0.007	0.009	-0.062	-0.057
	(0.687)	(0.670)	(0.021)	(0.021)	(0.112)	(0.114)
Observations	3,644	3,644	3,627	3,627	3,030	3,030
R-squared	0.440	0.455	0.465	0.471	0.144	0.166
HH char.	N	Y	N	Y	N	Y
Mean of non-exposed	156.88		5.840		-0.001	
group	(8.39)		(0.273)		(1.005)	

Notes: The maximum of 3 measures of lung capacity is used (in log). Test score is the the fraction correct response on both the Raven and mathematics tests. Covariates include male child, mother's age at delivery, and mother's education, urban residence, household size, household per capita expenditure (log), an asset index and a housing quality index. Mother's primary education takes the value one if she has primary education or lower. Per capita expenditure is in 2000 Rupiah (1 USD˜Rp. 10,000 in 2000). The regressions include district, children's year and month of birth fixed effects. All standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Potential channels

	(1)	(2)	(3)
Panel A. Maternal health			
	Mother's lung capacity	Mother's respiratory problem	Mother's BMI
Pregnant x affected	-0.0566*	0.0722*	0.270
	(0.0300)	(0.0428)	(0.334)
Pregnant	-0.0316**	0.0393	0.997***
	(0.0126)	(0.0253)	(0.197)
Observations	4,065	4,329	4,113
R-squared	0.333	0.131	0.213
Mean of non-exposed	5.665	0.199	22.516
group	(0.219)	(0.399)	(3.773)
Panel B. Contemporaneous effects on children			
	Any respiratory problem	Wasting	Birth weight (kg)
In-utero			-0.018
			(0.101)
0-12mo	0.180**	0.015	
	(0.070)	(0.047)	
13-24mo	0.206**	0.077	
	(0.080)	(0.053)	
Observations	1,539	1,254	3,305
R-squared	0.147	0.172	0.119
Mean of non-exposed	0.438	-0.442	3.169
group	(0.496)	(1.280)	(0.559)
Panel C. Parental response			
	Breast-fed>6mo	Birth spacing (months)	
In-utero	-0.006	2.621	
	(0.027)	(4.010)	
0-12mo	-0.004	4.394	
	(0.021)	(4.734)	
13-24mo	0.002	4.799	
	(0.024)	(5.940)	
Observations	4,500	8,485	
R-squared	0.121	0.155	
Mean of non-exposed	0.750	59.75	
group	(0.433)	(38.71)	

Notes: Panel A includes ever married reproductive aged women between 15 and 49 years who were surveyed in the second wave of the IFLS in 1997. Panel B, cols. 1-2 include children who were born between 1995 and the time of the survey in 1997 (children in this sample are between 0 and 30 months at the time of the survey). Panel B, col. 3 includes children who were born between 1995 and 2000. Panel C includes children who were born between 1995 and 2000. All analyses include district fixed effects. All standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01

Table 7: Pre-conception exposure to pollution in affected and unaffected areas

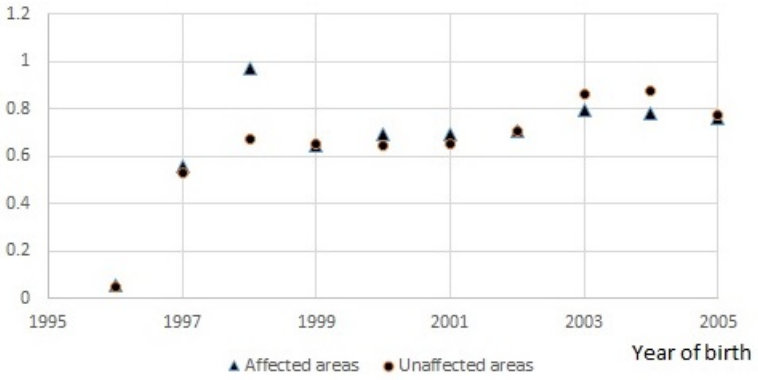
	(1) Height for age (2000)	(2) Height	(3) Lung capacity (2007)	(4) Test score	(5) Height	(6) Lung capacity (2014)	(7) Test score
Exposed in 6mo. pre-conception	-0.246	-0.352	-0.031	-0.016	-1.048	0.002	0.038
	(0.231)	(0.610)	(0.043)	(0.095)	(1.062)	(0.022)	(0.166)
In-utero	-0.327*	-1.586**	-0.054*	-0.051	-1.499*	0.035	0.100
	(0.167)	(0.601)	(0.030)	(0.127)	(0.783)	(0.025)	(0.123)
0-12mo	-0.417***	-0.699	-0.053**	-0.008	-0.486	0.025	-0.007
	(0.108)	(0.744)	(0.023)	(0.081)	(0.783)	(0.020)	(0.091)
13-24mo	-0.282***	-0.197	-0.056***	0.111	-0.872	0.009	-0.053
	(0.093)	(0.708)	(0.021)	(0.099)	(0.674)	(0.021)	(0.112)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.219	0.682	0.381	0.192	0.455	0.471	0.166

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the community level (enumeration area). * p<0.1, ** p<0.05, *** p<0.01.

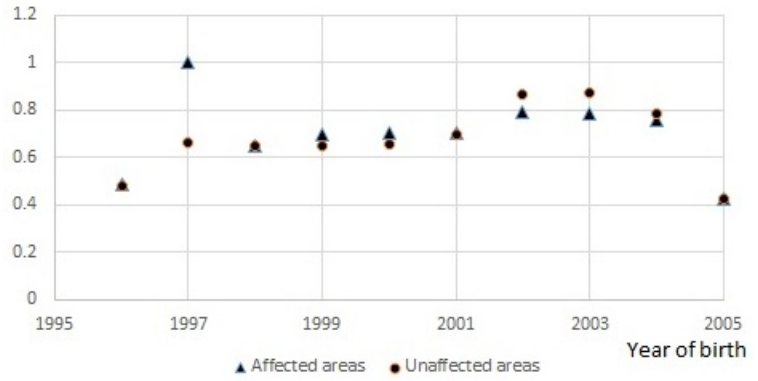
A Appendix: Supplementary Figures and Tables

Figure A.1: Average aerosol index in early life

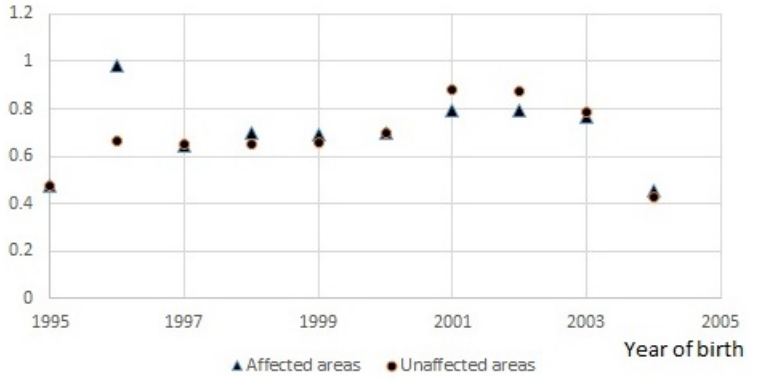
Panel A. Average aerosol index in utero



Panel B. Average aerosol index at 0-12 months



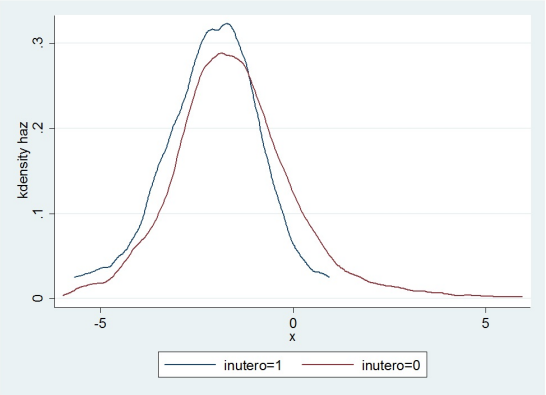
Panel C. Average aerosol index at 13-24 months



Notes: Affected takes the value 1 if the child's district of birth had an aerosol index of more than 1.5 for at least 3 days in a month. The aerosol index ranges from 0 to almost 6 in our sample.

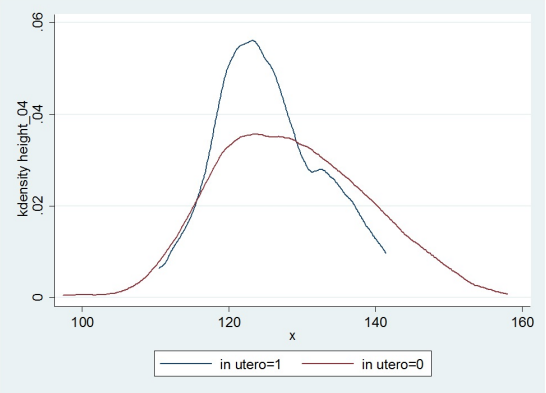
Figure A.2: Distribution of children's height

Panel A. Height for age z-score 3 years post exposure



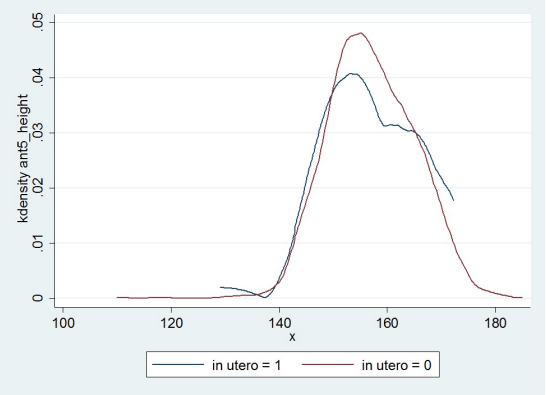
Kolmogorov-Smirnov test p-value 0.006.

Panel B. Height in 2007: 10 years post exposure



Kolmogorov-Smirnov test p-value 0.083.

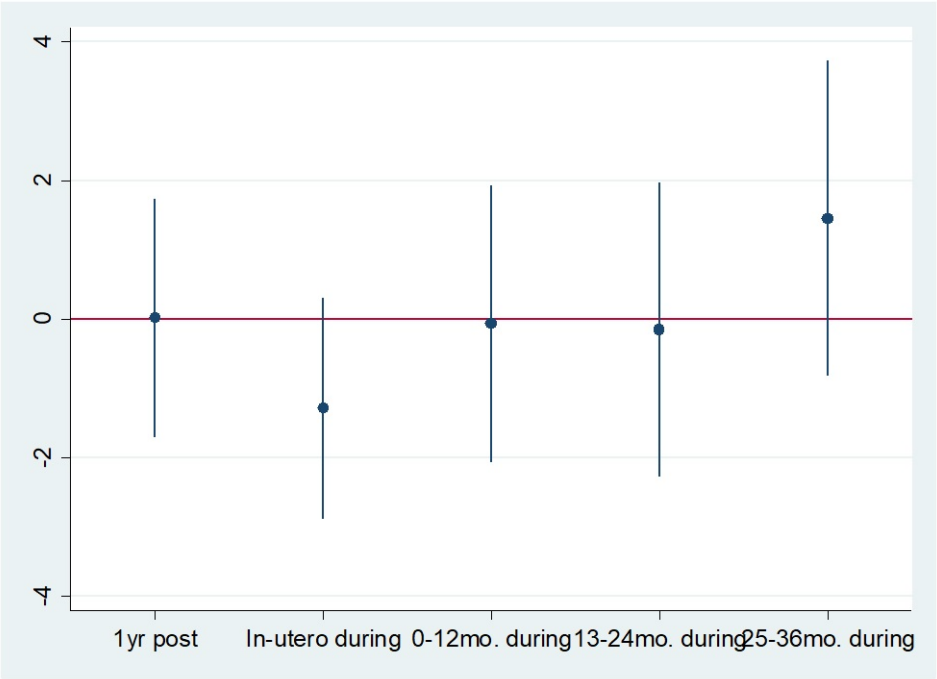
Panel C. Height in 2014: 17 years post exposure



Kolmogorov-Smirnov test p-value 0.360.

Notes: In utero takes the value 1 if the child's district of birth had an aerosol index of more than 1.5 for at least 3 days in a month. The aerosol index ranges from 0 to almost 6 in our sample.

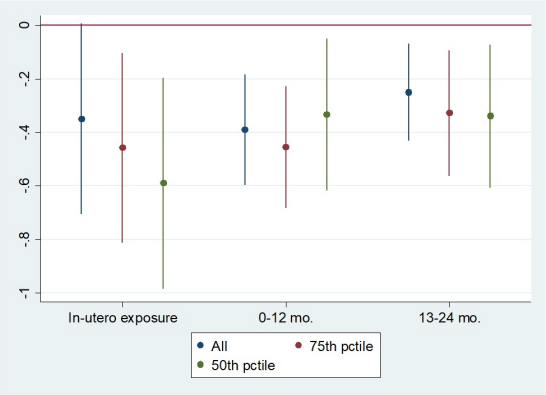
Figure A.3: The effects on height in 2014 for the affected and non-affected cohorts



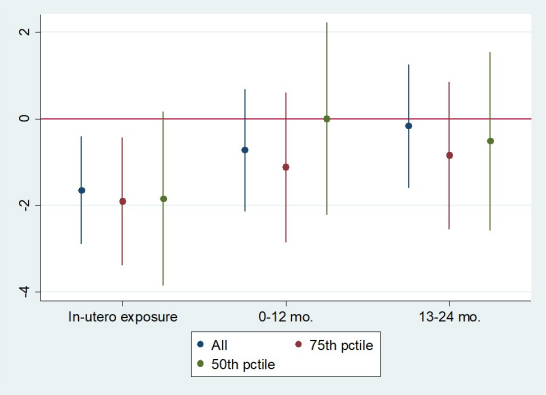
Notes: Coefficient of the interaction between born in areas affected by the fires and month-year of birth. The omitted category is children conceived 2 years after the fires. All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area).

Figure A.4: Robustness: Alternative sample by distance from district of birth to TOMS grid (pollution measure)

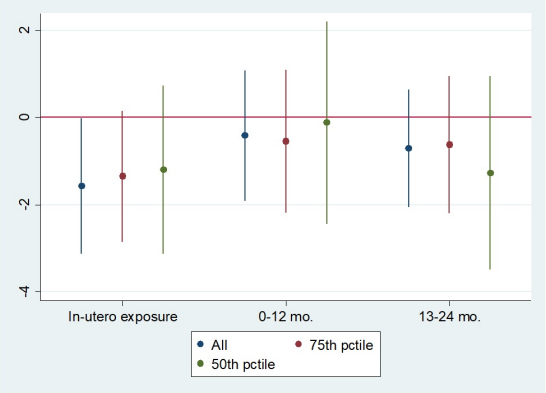
Panel A. Height for age z-score (in 2000)



Panel B. Height 10 years post exposure (in 2007)



Panel C. Height 17 years post exposure (in 2014)



Notes: The average distance between the district of birth centroid to the TOMS grid is 29 kilometers, with a maximum of 50 kilometers. The 75th percentile corresponds to a distance of 38 kilometers between the district of birth centroid and TOMS grid (pollution measure). The 50th percentile corresponds to a distance of 32 kilometers between the district of birth centroid and TOMS grid (pollution measure). The regressions include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area).

Table A.1: Association between burnt area and aerosol index

	TOMS aerosol index	Number of days the aerosol index is greater than 1.5 in a month
Monthly burnt area in hectares	0.00004**	0.00030**
	(0.00002)	(0.00012)
Observations	6,475	6,475
R-squared	0.46	0.46

Notes: The sample includes information from TOMS grids from January 1997 to December 1997. Monthly burnt area comes from the Global Fire Emissions Database (GFED4). Robust standard errors in parentheses clustered at the pollution grid level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Restricted sample: Children with measured height in every wave

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Height for age z-score (2000)	Height	Lung capacity (2007)	Cognitive test	Height	Lung capacity (2014)	Cognitive test
In-utero	-0.468*	-2.041**	-0.120**	-0.115	-1.428	-0.006	0.016
	(0.251)	(0.957)	(0.057)	(0.171)	(0.982)	(0.032)	(0.160)
0-12mo	-0.410***	-0.250	-0.068	0.027	-0.442	0.028	-0.016
	(0.134)	(0.931)	(0.051)	(0.148)	(0.761)	(0.026)	(0.099)
13-24mo	-0.299**	-0.005	-0.097	0.032	-0.814	-0.007	-0.139
	(0.131)	(1.023)	(0.059)	(0.191)	(1.060)	(0.028)	(0.153)
Observations	2,440	2,504	1,302	1,984	2,506	2,493	2,100
R-squared	0.234	0.647	0.415	0.212	0.466	0.477	0.199
Mean of non-exposed group	-1.464	0.378	127.88	5.339	156.88	5.840	-0.001
	(2.291)	(0.485)	(11.89)	(0.290)	(8.39)	(0.273)	(1.005)

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table A.3: Alternative clustering: community level (enumeration area)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Height for age z-score (2000)	Height	Lung capacity (2007)	Cognitive test	Height	Lung capacity (2014)	Cognitive test
In-utero	-0.352**	-1.628**	-0.056**	-0.052	-1.559*	0.036*	0.102
	(0.167)	(0.758)	(0.028)	(0.116)	(0.827)	(0.021)	(0.125)
0-12mo	-0.380***	-0.653	-0.050**	-0.004	-0.381	0.024	-0.013
	(0.114)	(0.839)	(0.025)	(0.103)	(0.697)	(0.023)	(0.091)
13-24mo	-0.251**	-0.176	-0.054**	0.111	-0.705	0.010	-0.057
	(0.110)	(0.739)	(0.023)	(0.095)	(0.672)	(0.022)	(0.118)
Observations	3,796	4,184	2,390	3,348	3,648	3,631	3,034
R-squared	0.219	0.682	0.379	0.192	0.457	0.472	0.167
Mean of non-exposed	-1.464	0.378	127.88	5.339	156.88	5.840	-0.001
group	(2.291)	(0.485)	(11.89)	(0.290)	(8.39)	(0.273)	(1.005)

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the community level (enumeration area). * p<0.1, ** p<0.05, *** p<0.01.

Table A.4: Pre-trends for children's height before the 1997 fires

	(1)
	Height for age z-score
Birth year 1988 x affected	0.156
	(0.194)
Birth year 1989 x affected	0.093
	(0.166)
Birth year 1990 x affected	0.014
	(0.205)
Birth year 1991 x affected	0.147
	(0.181)
Birth year 1992 x affected	-0.163
	(0.159)
Birth year 1993 x affected	-0.006
	(0.180)
Observations	3,687
R-squared	0.206

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Absence of effects in non-treated cohorts

	(1) Height	(2) Lung capacity	(3) Cognitive test	(4) Height	(5) Lung capacity	(6) Cognitive test
		2007			2014	
Conceived 1 year after	-0.139	0.010	0.129	0.018	0.039	-0.013
	(0.498)	(0.050)	(0.142)	(0.857)	(0.027)	(0.021)
In-utero during	-1.231*	-0.063	0.024	-1.279	0.069**	0.015
	(0.717)	(0.039)	(0.151)	(0.796)	(0.030)	(0.024)
0-12 mo. during	-0.876	-0.069***	0.069	-0.060	0.058**	0.008
	(0.725)	(0.020)	(0.129)	(0.996)	(0.024)	(0.023)
13-24 mo. during	0.533	-0.056**	0.184	-0.147	0.058**	-0.014
	(0.885)	(0.024)	(0.161)	(1.059)	(0.026)	(0.022)
25-36 mo. during	1.834*		0.119	1.459	0.049*	0.027
	(1.041)		(0.139)	(1.136)	(0.025)	(0.026)
Observations	4,181	2,389	3,414	3,645	3,628	3,032
R-squared	0.683	0.380	0.187	0.457	0.473	0.171

Notes: We estimate the following equation: $y_{idmy} = \Sigma_{y=age\_2}^{y=age\_-2} Affected\_Dis_d * cohort_y + \beta X_i + \tau_y + \mu_m + \rho_d + \epsilon_{idmy}$, where our coefficients of interest are the interactions between being born in a district affected by the 1997 forest fires ($Affected\_Dis_d$) and cohort: 25-36 months during the fires, 13-24 months during the fires, 0-12 months during the fires, in-utero during the fires, or conceived 1 year after. The omitted cohort is the one conceived two years or more after the fires (born in late 1999 and 2000). However, for lung capacity in 2007, the omitted category is the cohort age 3 during the fires since that cohort was not eligible for the lung capacity measures (younger than age 9). All analyses include household characteristics, as well as district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Falsification test: "Fake fires date"

	(1)	(2)
	Height	Lung capacity
In-utero	0.786	0.012
	(0.604)	(0.015)
0-12mo	0.104	0.000
	(0.777)	(0.021)
13-24mo	0.616	0.034*
	(0.598)	(0.018)
Observations	4,542	3,759
R-squared	0.674	0.381

Notes: Children's outcomes in 2014 for those born between 2001 and 2006 with 'exposure' in-utero to 13-24 months. All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). *** p<0.01, ** p<0.05, * p<0.1

Table A.7: Robustness: The effect of household unemployment, income loss, and crop loss in 1997

	(1) Height for age z-score (2000)	(2) Height	(3) Lung capacity (2007)	(4) Cognitive test	(5) Height	(6) Lung capacity (2014)	(7) Cognitive test
In-utero	-0.347*	-1.816***	-0.059*	-0.036	-1.564*	0.036	0.105
	(0.180)	(0.677)	(0.034)	(0.131)	(0.810)	(0.026)	(0.123)
0-12mo	-0.380***	-0.486	-0.051**	-0.049	-0.262	0.028	-0.022
	(0.108)	(0.744)	(0.022)	(0.073)	(0.734)	(0.019)	(0.088)
13-24mo	-0.256***	-0.245	-0.050**	0.097	-0.675	0.006	-0.080
	(0.088)	(0.754)	(0.019)	(0.095)	(0.652)	(0.021)	(0.117)
Province unemployment	-0.013	-0.120	-0.002	-0.006	-0.032	-0.008***	0.017
	(0.020)	(0.137)	(0.005)	(0.011)	(0.096)	(0.003)	(0.018)
Crop	0.122	0.366	0.068**	0.088	-0.250	0.049**	0.211**
loss	(0.115)	(0.723)	(0.028)	(0.107)	(0.530)	(0.021)	(0.101)
Income	-0.154	-0.716	0.004	0.007	0.072	-0.009	0.170*
loss	(0.124)	(0.954)	(0.018)	(0.081)	(0.839)	(0.022)	(0.102)
Observations	3,790	3,991	2,273	3,200	3,609	3,592	3,000
R-squared	0.219	0.683	0.389	0.194	0.458	0.475	0.171

Notes: Province unemployment is the change in province of birth unemployment at the time of birth. All analyses include household characteristics, as well as district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table A.8: Selective fertility and migration

	(1)	(2)	(3)
	Pregnant during the fires	Pregnant after the fires	Moved after 1997
Affected district	0.0345	-0.0259	-0.0378
	(0.0506)	(0.0380)	(0.0894)
Affected x Low education	0.0130	-0.0128	0.0282
	(0.0231)	(0.0268)	(0.0240)
Affected x Urban	-0.0455	-0.00688	-0.0079
	(0.0327)	(0.0199)	(0.041)
Affected x Poor	-0.0307	-0.0419*	0.0245
	(0.0298)	(0.0226)	(0.0187)
Affected x low asset index	-0.0328	0.0395	0.0157
	(0.0252)	(0.0308)	(0.0214)
Affected x maternal age	-0.0129	-0.0456	-0.0128
< 20	(0.0789)	(0.0372)	(0.0231)
Affected x maternal age	-0.00812	0.0313	0.0193
21-25	(0.0426)	(0.0347)	(0.0242)
Affected x maternal age	-0.0109	-0.0141	-0.0053
26-30	(0.0297)	(0.0320)	(0.0263)
Observations	4,338	5,425	5,875
R-squared	0.062	0.038	0.119

Notes: Mother's primary education takes the value one if she has primary education or lower. Per capita expenditure is in 2000 Rupiah (1 USD˜Rp. 10,000 in 2000). Poor takes the value one if the household per capita expenditure is at the bottom two quintiles. Low asset index takes the value one if the household is at the bottom two quintiles of the index. Province fixed effects included. All standard errors are clustered at the grid (pollution measure area). * p<0.1, ** p<0.05, *** p<0.01.

Online Appendix - Not for Publication

B Heterogeneous treatment effects

Exposure to air pollution during the fires may affect children with different backgrounds differently, so we explore heterogeneous treatment effects. We explore heterogeneity by gender, followed by socio-economic characteristics. We explore the following heterogeneities: maternal education, urban residence, poverty (defined as the bottom two quintiles of per capita expenditure), agricultural participation, and chronic indoor pollution from wood-burning stove.

We begin with gender since the fragile male hypothesis argues that males are more vulnerable to shocks in early-life compared to females (Kraemer, 2000). According to this hypothesis, males would be more affected by early-life exposure to air pollution. Table B.1 presents the short-term (column 1) and longer-term effects of exposure to air pollution on males and females (columns 2-4 correspond to outcomes 10 years post exposure, columns 5-7 correspond to outcomes 17 years post exposure). Among children who were exposed to the pollution in-utero, we find no evidence of heterogeneity by gender in the short and longer-term. This is consistent with Jayachandran (2009) who found no evidence of gender selective mortality immediately after the fires. Among children who were exposed to the fires in their first or second year of life, we find mixed results– males appear to be adversely affected for some outcomes, but positively affected for others.

To explore the role of socio-economic status, we analyze heterogeneity by mother's education in Table B.2. The variable low education takes the value 1 if the child's mother has primary education or lower. We find no significant heterogeneous effects by mother's education among children who were exposed to air pollution in-utero or in their first year of life, even though higher maternal education is generally associated with protective health effects. Among children who were exposed to air pollution in their second year of life, we find that low maternal education is associated with lower height and lower lung capacity 10 years post exposure, but the effects are no longer significant 17 years post exposure. The results suggest that maternal education has only some mediating effects.

We also examine heterogeneity by urban residence, which is typically associated with protective health effects (Fink, Günther and Hill, 2014). In our setting, urban residents generally have better access to infrastructure and health care; therefore, we expect urban residents to be able to better engage in mitigating behavior. We find little heterogeneity by urban residence (Table B.3), suggesting that children in urban and rural areas were equally affected. We find no heterogeneous treatment effects by urban residence among children who

were exposed to the fires in-utero and mixed results for children in urban areas who were exposed in their first or second year of life. In spite of the disparity in urban and rural access to health care services, children residing in rural areas who were exposed to the fires do not seem to be more adversely affected by the fires.

Similarly, we also find no heterogeneous treatment effects by poverty, defined as the bottom two quintiles of per capita expenditure (Table B.4). This corresponds to an average per capita expenditure of 98,000 *Rupiah* per month (in 2000 *Rupiah*, approximately USD 10).[33] A closely related issue concerns households engaged in agriculture, since the drought that was associated with the El Niño phenomenon might have affected these households more severely than non-agricultural households. We find limited evidence of heterogeneity by agricultural status for children exposed to the fires in early-life (Table B.5). These results highlight the indiscriminate health effects of early-life exposure to acute air pollution. Further research on air pollution and avoidance behavior in developing countries will allow us to explore the role of information and socio-economic status in mitigating the effects of pollution on long-term human capital outcomes (Bharadwaj et al., 2017).

We also explore differential effects by chronic exposure to indoor air pollution. We include an interaction term between outdoor air pollution from the fires and indoor pollution from wood stove– a major source of indoor air pollution in many developing countries (table B.6). We find that children who were exposed to the smoke from wood burning stoves at home are likely to have poorer health outcomes and test scores. However, we find no significant heterogeneous effects on children's height or lung capacity by exposure to indoor air pollution from wood burning stoves. This is consistent with earlier results that suggest that the effects of air pollution on early-life mortality can be attributed to the fires, not exposure to indoor air pollution from wood burning stoves (Jayachandran, 2009).

[33] When we use the lowest quintile of per capita expenditure, we find worse outcomes in height for exposed children. This is consistent with earlier work by Jayachandran (2009) that finds larger effects in poorer communities.

Table B.1: Heterogeneous treatment effects by gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age	Height	Lung capacity	Test score	Height	Lung capacity	Test score
Male	-0.079*	-0.691***	0.085***	-0.017	8.926***	0.311***	-0.090**
	(0.045)	(0.202)	(0.015)	(0.022)	(0.267)	(0.009)	(0.045)
In-utero	-0.283	-2.141**	-0.074	-0.171	-1.835*	0.021	0.166
	(0.191)	(0.880)	(0.049)	(0.128)	(0.932)	(0.030)	(0.136)
0-12mo	-0.572***	-1.142	-0.045	0.199	-1.701*	0.025	0.114
	(0.155)	(1.144)	(0.035)	(0.130)	(0.941)	(0.021)	(0.130)
13-24mo	-0.256**	1.766**	-0.040	0.110	-1.395*	-0.036	-0.015
	(0.117)	(0.761)	(0.024)	(0.122)	(0.702)	(0.027)	(0.132)
In-utero x Male	-0.147	0.902	0.028	0.226	0.487	0.028	-0.125
	(0.317)	(1.378)	(0.056)	(0.142)	(1.122)	(0.041)	(0.179)
0-12mo x Male	0.351	0.839	-0.008	-0.359**	2.596**	-0.002	-0.229
	(0.239)	(1.593)	(0.041)	(0.166)	(1.151)	(0.033)	(0.216)
13-24mo x Male	0.015	-3.776***	-0.025	0.003	1.410	0.099***	-0.087
	(0.184)	(1.130)	(0.031)	(0.142)	(0.958)	(0.032)	(0.154)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.219	0.683	0.381	0.193	0.456	0.472	0.166

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table B.2: Heterogeneous treatment effects by mother's education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age	Height	Lung capacity	Test score	Height	Lung capacity	Test score
Mother's education: primary	-0.134**	-0.483**	0.003	-0.238***	-0.189	-0.019***	-0.234***
	(0.060)	(0.221)	(0.016)	(0.050)	(0.254)	(0.007)	(0.026)
In-utero	-0.407*	-1.961**	-0.082*	-0.075	-2.010*	0.019	-0.036
	(0.223)	(0.796)	(0.048)	(0.161)	(1.182)	(0.035)	(0.160)
0-12mo	-0.337***	-0.657	-0.030	0.036	-0.155	0.032	0.042
	(0.114)	(0.826)	(0.028)	(0.092)	(0.924)	(0.024)	(0.126)
13-24mo	-0.165	0.935	-0.021	0.151	-0.654	0.030	-0.174
	(0.130)	(0.954)	(0.026)	(0.113)	(1.000)	(0.024)	(0.128)
In-utero x Low maternal education	0.124	0.807	0.060	0.060	1.064	0.039	0.316*
	(0.271)	(1.214)	(0.053)	(0.227)	(1.440)	(0.048)	(0.178)
0-12mo x Low maternal education	-0.140	-0.239	-0.054	-0.118	-0.535	-0.021	-0.116
	(0.186)	(1.151)	(0.036)	(0.149)	(1.100)	(0.035)	(0.171)
13-24mo x Low maternal education	-0.187	-2.192**	-0.067**	-0.087	-0.219	-0.046	0.245
	(0.205)	(0.955)	(0.029)	(0.205)	(1.430)	(0.030)	(0.218)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.219	0.682	0.383	0.192	0.455	0.472	0.167

Notes: Low education takes the value one if mother's education is primary or lower. All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table B.3: Heterogeneous treatment effects by urban residence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age	Height	Test score	Lung capacity	Height	Lung capacity	Test score
Urban	0.302	0.100	0.002	-0.056	0.408	-0.004	0.109*
	(0.186)	(0.519)	(0.024)	(0.075)	(0.453)	(0.012)	(0.061)
In-utero	-0.475**	-1.554**	-0.038	-0.175	-1.378**	0.043	0.090
	(0.210)	(0.713)	(0.035)	(0.142)	(0.660)	(0.029)	(0.136)
0-12mo	-0.387***	-0.873	-0.045	-0.125	0.216	0.031	-0.029
	(0.129)	(0.922)	(0.028)	(0.095)	(0.927)	(0.025)	(0.091)
13-24mo	-0.158	-0.310	-0.052*	0.052	0.113	-0.004	-0.016
	(0.132)	(1.091)	(0.027)	(0.144)	(0.974)	(0.031)	(0.123)
In-utero x Urban	0.367	-0.183	-0.056	0.369	-0.639	-0.023	0.065
	(0.247)	(1.344)	(0.052)	(0.233)	(1.440)	(0.035)	(0.230)
0-12mo x Urban	-0.021	0.648	-0.013	0.355*	-1.693	-0.018	0.059
	(0.157)	(1.284)	(0.042)	(0.209)	(1.357)	(0.030)	(0.128)
13-24mo x Urban	-0.223	0.392	-0.004	0.169	-1.962*	0.027	-0.085
	(0.200)	(1.322)	(0.046)	(0.183)	(1.170)	(0.040)	(0.178)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.220	0.682	0.381	0.194	0.456	0.471	0.166

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table B.4: Heterogeneous treatment effects by poverty

Panel A: Per capita expenditure Quintile 1 and 2 defined as poor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age z-score	Height	Lung capacity	Cognitive test	Height	Lung capacity	Cognitive test
Poor household	-0.061	-0.016	-0.019	-0.046	0.131	-0.001	-0.121*
	(0.093)	(0.545)	(0.017)	(0.066)	(0.376)	(0.014)	(0.070)
In-utero	-0.183	-1.117*	-0.068	-0.082	-0.670	0.040	0.042
	(0.181)	(0.562)	(0.042)	(0.139)	(0.693)	(0.030)	(0.145)
0-12mo	-0.403**	-0.219	-0.030	-0.045	-0.465	0.020	0.022
	(0.159)	(0.895)	(0.028)	(0.114)	(0.999)	(0.030)	(0.102)
13-24mo	-0.328***	0.069	-0.049**	0.115	-1.370	0.014	-0.117
	(0.118)	(0.946)	(0.022)	(0.138)	(0.937)	(0.023)	(0.108)
In-utero x Poor	-0.367	-1.106	0.027	0.058	-2.339	-0.013	0.159
	(0.318)	(1.124)	(0.055)	(0.197)	(1.808)	(0.039)	(0.180)
0-12mo x Poor	0.060	-0.940	-0.047	0.091	0.371	0.011	-0.079
	(0.174)	(1.119)	(0.031)	(0.164)	(1.331)	(0.047)	(0.139)
13-24mo x Poor	0.211	-0.508	-0.009	-0.004	1.711	-0.013	0.130
	(0.239)	(1.697)	(0.032)	(0.221)	(1.840)	(0.039)	(0.172)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.219	0.682	0.382	0.192	0.456	0.471	0.167

Table B.4: Cont. Heterogeneous treatment effects by poverty

Panel B: Per capita expenditure Quintile 1 defined as poorest

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age z-score	Height	Lung capacity	Cognitive test	Height	Lung capacity	Cognitive test
Poorest	0.163**	0.864**	0.006	0.033	0.444	0.002	-0.020
	(0.079)	(0.371)	(0.015)	(0.068)	(0.406)	(0.013)	(0.057)
In-utero	-0.228	-1.131**	-0.046	-0.073	-1.279	0.043	0.042
	(0.140)	(0.519)	(0.037)	(0.132)	(0.837)	(0.027)	(0.124)
0-12mo	-0.369***	-0.003	-0.055**	0.034	-0.341	0.020	0.025
	(0.122)	(0.735)	(0.023)	(0.093)	(0.802)	(0.022)	(0.084)
13-24mo	-0.235**	0.007	-0.050**	0.171	-0.759	0.019	-0.102
	(0.091)	(0.753)	(0.020)	(0.106)	(0.697)	(0.022)	(0.101)
In-utero x Poorest	-0.734	-1.937*	-0.057	0.132	-1.889	-0.051	0.391*
	(0.541)	(1.109)	(0.065)	(0.258)	(2.608)	(0.051)	(0.208)
0-12mo x Poorest	-0.073	-3.160*	0.037	-0.261	0.017	0.030	-0.215
	(0.217)	(1.781)	(0.058)	(0.185)	(1.510)	(0.038)	(0.221)
13-24mo x Poorest	-0.121	-0.726	-0.015	-0.346	0.171	-0.057	0.251
	(0.208)	(1.184)	(0.050)	(0.214)	(1.731)	(0.048)	(0.281)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.220	0.683	0.381	0.193	0.455	0.472	0.166

Notes: Poverty defined as the bottom two quintiles of per capita expenditure. Poorest defined as the lowest quintile of per capita expenditure. All analyses include household characteristics, as well as district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table B.5: Heterogeneous treatment effects by agricultural participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age	Height	Test score	Lung capacity	Height	Lung capacity	Test score
Agricultural	-0.057	0.060	-0.009	0.003	0.763*	0.015	-0.142
household	(0.096)	(0.414)	(0.015)	(0.089)	(0.407)	(0.018)	(0.095)
In-utero	-0.358**	-1.908***	-0.065*	-0.048	-1.837**	0.034	0.106
	(0.176)	(0.624)	(0.036)	(0.128)	(0.839)	(0.027)	(0.130)
0-12mo	-0.428***	-0.585	-0.056**	-0.005	-0.843	0.022	-0.045
	(0.114)	(0.795)	(0.024)	(0.099)	(0.733)	(0.021)	(0.101)
13-24mo	-0.349***	-0.428	-0.054**	0.119	-1.053	0.019	-0.143
	(0.099)	(0.760)	(0.021)	(0.105)	(0.715)	(0.022)	(0.120)
In-utero x Agri	0.112	5.886	0.141*	-0.324	5.412	0.037	-0.391*
	(0.583)	(4.211)	(0.075)	(0.251)	(4.024)	(0.076)	(0.204)
0-12mo x Agri	0.221	-0.564	0.021	0.002	1.881	0.009	0.206
	(0.195)	(1.221)	(0.029)	(0.236)	(1.922)	(0.041)	(0.194)
13-24mo x Agri	0.362	1.010	0.006	-0.022	1.081	-0.035	0.334
	(0.263)	(1.077)	(0.036)	(0.208)	(1.131)	(0.048)	(0.219)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.219	0.682	0.381	0.192	0.457	0.472	0.167

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table B.6: Heterogeneous treatment effects by use of wood stove

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age z-score	Height	Lung capacity	Cognitive test	Height	Lung capacity	Cognitive test
Wood stove	-0.147**	-1.117***	-0.034	-0.151***	-0.885***	-0.012	-0.134*
	(0.069)	(0.370)	(0.021)	(0.041)	(0.315)	(0.012)	(0.071)
In-utero	-0.281	-1.187	-0.088***	0.127	-1.730*	0.046	0.132
	(0.171)	(0.846)	(0.031)	(0.176)	(0.972)	(0.032)	(0.122)
0-12mo	-0.398***	-1.294**	-0.062**	0.003	-1.046	0.031	-0.065
	(0.104)	(0.641)	(0.026)	(0.101)	(0.811)	(0.020)	(0.113)
13-24mo	-0.364***	0.183	-0.059***	0.142	-0.842	0.003	-0.096
	(0.111)	(0.828)	(0.021)	(0.088)	(0.580)	(0.021)	(0.133)
In-utero x Wood stove	-0.220	-0.843	0.070	-0.358*	0.664	-0.034	-0.061
	(0.289)	(1.303)	(0.054)	(0.205)	(1.337)	(0.055)	(0.223)
0-12mo x Wood stove	0.034	1.422	0.026	-0.038	2.595*	-0.022	0.186
	(0.171)	(1.163)	(0.041)	(0.104)	(1.520)	(0.032)	(0.209)
13-24mo x Wood stove	0.296	-0.726	0.019	-0.051	0.229	0.026	0.137
	(0.251)	(1.265)	(0.032)	(0.173)	(1.320)	(0.048)	(0.177)
Observations	3,790	4,178	2,385	3,342	3,643	3,626	3,030
R-squared	0.220	0.683	0.383	0.196	0.456	0.472	0.167

Notes: All analyses include household characteristics, as well as district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

C Alternative measures

We examine robustness of our results to alternative definitions of the shock, and explore different thresholds of the aerosol index for sensitivity. First, we re-define the treatment variables in equation 1 using a higher threshold of the aerosol index (more than 1.75) and a lower threshold (more than 1.25) for 3 days or more in a month. Second, we explore a continuous measure, defined as the mean aerosol index in-utero, during 0-12 months and 13-24 months. Third, we explore the presence of non-linearities by re-defining alternative treatment variables in two ways: by varying the number of days exposed and by varying the levels.

First, we re-define the treatment variables in equation 1 using a higher threshold of the aerosol index (more than 1.75) and a lower threshold (more than 1.25) for 3 days or more in a month. At a higher threshold of 1.75 (Table C.1, Panel A), we find persistent effects on height among children who were exposed to the fires in-utero at 3, and 10 years post-exposure (although only 4.5 percent of the sample were exposed to the higher pollution threshold). The negative effects of in-utero exposure on height 17 years after the shock are negative and similar in magnitude to our main specification, albeit they are not statistically significant. We also find that children's respiratory function is affected 10 years after exposure, but the results are not significant 17 years after exposure, which mimics the findings from our main specification. We find no significant effects on cognitive test 10 or 17 years after exposure. Using a lower threshold of aerosol index of 1.25 (Panel B), we find qualitatively similar results to our earlier estimates.[34] Exposure to an aerosol index of 1.25 for at least 3 days in a month in early-life continues to be associated with persistent health effects. Children who were exposed to a lower pollution threshold in-utero are about 1.8 centimeters shorter 10 years post exposure, which is similar to our earlier finding. The negative effects of in-utero exposure on height 17 years after the shock are negative and similar in magnitude to our main specification (1.65 centimeters). According to these robustness checks, the health effects persist even when children are exposed to lower pollution levels, suggesting the importance of the longer-term effects of pollution.

We also explore a continuous measure, defined as the mean aerosol index in-utero, during 0-12 months and 13-24 months. The results suggest that a 1 unit increase in the aerosol index in early-life is associated with some adverse later life outcomes, but the results may mask non-linearities in the effect of air pollution (Table C.2, Panel A). Next, we use two alternative measures to explore non-linearities. 1) We use aerosol index above 1.5 as the

[34] When we use an even lower threshold, an aerosol index of 1.00, the results are qualitatively similar, but most of the effects are not statistically significant.

cutoff and vary the threshold of exposure at 0 days (reference category), 1-2 days, 3-6 days, and 7 days or more (Table C.3). 2) We use 3 days or more of exposure as the cutoff and vary the aerosol index threshold at 0 to 1 (reference category), 1 to 1.5, 1.5 to 2.25 and above 2.25 (Table C.4). These alternative measures provide evidence of the non-linear effects of exposure to air pollution, and worse later life outcomes appear to be driven by exposure to higher pollution days as measured by the aerosol index.

Table C.1: Alternative pollution measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age z-score	Height	Lung capacity	Cognitive test	Height	Lung capacity	Cognitive test
Panel A. Aerosol index > 1.75							
In-utero	-0.303*	-1.236**	-0.073**	-0.040	-1.214	0.037	0.064
	(0.176)	(0.465)	(0.035)	(0.139)	(0.811)	(0.027)	(0.132)
0-12mo	-0.364***	-0.535	-0.044*	0.020	-0.127	0.034*	-0.011
	(0.103)	(0.761)	(0.024)	(0.088)	(0.839)	(0.020)	(0.093)
13-24mo	-0.201*	-0.012	-0.048**	0.105	-0.941	0.010	-0.029
	(0.105)	(0.617)	(0.019)	(0.117)	(0.728)	(0.024)	(0.119)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.218	0.681	0.381	0.192	0.455	0.471	0.165
Panel B. Aerosol index > 1.25							
In-utero	-0.359**	-1.787***	-0.058*	-0.070	-1.650**	0.019	0.023
	(0.157)	(0.634)	(0.032)	(0.121)	(0.735)	(0.024)	(0.117)
0-12mo	-0.297***	-0.239	-0.044**	-0.003	0.015	0.031*	-0.055
	(0.102)	(0.692)	(0.021)	(0.075)	(0.714)	(0.018)	(0.084)
13-24mo	-0.210**	-0.185	-0.049***	0.097	-0.728	0.009	-0.050
	(0.092)	(0.706)	(0.018)	(0.094)	(0.638)	(0.020)	(0.108)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.218	0.682	0.381	0.192	0.455	0.471	0.165

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table C.2: Continuous pollution measure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age z-score	Height	Lung capacity	Cognitive test	Height	Lung capacity	Cognitive test
Panel A. Mean aerosol index							
In-utero	-0.161	-1.184	-0.055	0.003	-0.541	0.007	0.015
	(0.180)	(0.927)	(0.039)	(0.129)	(0.898)	(0.028)	(0.032)
0-12mo	-0.807***	-1.837	-0.018	-0.297**	-0.748	0.016	-0.009
	(0.198)	(1.360)	(0.039)	(0.147)	(1.306)	(0.033)	(0.023)
13-24mo	-0.652***	-2.580**	-0.115***	-0.009	-0.470	-0.008	-0.001
	(0.161)	(1.124)	(0.040)	(0.160)	(1.060)	(0.030)	(0.033)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.220	0.682	0.381	0.192	0.454	0.471	0.169
Panel B. Number of days exposed to aerosol index above 1.5							
In-utero	-0.010	-0.074***	-0.002	-0.002	-0.050	0.001	0.001
	(0.008)	(0.025)	(0.002)	(0.005)	(0.030)	(0.001)	(0.001)
0-12mo	-0.010*	-0.032	-0.001	-0.003	-0.041	0.000	-0.000
	(0.005)	(0.026)	(0.001)	(0.003)	(0.026)	(0.001)	(0.001)
13-24mo	-0.004	0.007	-0.001	0.002	-0.010	-0.000	-0.000
	(0.005)	(0.023)	(0.001)	(0.003)	(0.027)	(0.001)	(0.001)
Observations	3,791	4,181	2,389	3,345	3,645	3,628	3,032
R-squared	0.216	0.682	0.378	0.191	0.456	0.471	0.170
Mean of non-exposed group	-1.478	127.91	5.339	0.008	156.89	5.841	-0.001
	(2.094)	(11.907)	(0.289)	(0.996)	(8.387)	(0.273)	(1.005)

Notes: The treatment variable in Panel A is the mean aerosol index in each stage of life. The mean aerosol index in-utero is 0.456 (0.332), age 1 is 0.565 (0.276) and age 2 is 0.682 (0.169). The treatment variable in Panel B is the number of days exposed to the mean aerosol index in each stage of life. All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table C.3: Alternative pollution measure: exposure to aerosol index of at least 1.5 for different numbers of days

	(1) 2000 Height for age z-score	(2)	(3) 2007 Lung capacity	(4)	(5)	(6) 2014 Lung capacity	(7)
		Height		Cognitive test	Height		Cognitive test
In-utero 1-2 days	0.090	-0.796	-0.063	-0.018	1.213*	-0.018	-0.017
	(0.178)	(0.560)	(0.040)	(0.079)	(0.668)	(0.020)	(0.019)
In-utero 3-6 days	-0.421*	-2.228***	-0.091***	-0.205*	-1.553*	-0.071**	-0.009
	(0.238)	(0.630)	(0.029)	(0.112)	(0.848)	(0.032)	(0.019)
In-utero 7 days	-0.372**	-1.500**	-0.058*	-0.035	-1.544**	0.018	-0.004
or more	(0.158)	(0.571)	(0.032)	(0.096)	(0.747)	(0.018)	(0.020)
0-12mo 1-2 days	0.024	0.716**	0.013	-0.002	0.132	0.014	-0.017
	(0.094)	(0.354)	(0.019)	(0.058)	(0.423)	(0.011)	(0.012)
0-12mo 3-6 days	-0.110	0.090	-0.044	0.092	-0.529	-0.007	0.004
	(0.164)	(0.411)	(0.034)	(0.085)	(0.635)	(0.016)	(0.016)
0-12 mo 7 days	-0.090	0.432	-0.038*	-0.078	-0.040	-0.001	-0.017
or more	(0.118)	(0.463)	(0.022)	(0.059)	(0.518)	(0.013)	(0.013)
13-24mo 1-2 days	0.028	-0.096	-0.012	0.049	0.353	-0.021	0.007
	(0.076)	(0.429)	(0.018)	(0.058)	(0.326)	(0.015)	(0.008)
13-24mo 3-6 days	0.200*	0.564	-0.009	-0.014	0.419	0.009	-0.001
	(0.113)	(0.562)	(0.030)	(0.088)	(0.729)	(0.018)	(0.012)
13-24 mo 7 days	-0.033	-0.016	-0.038**	-0.009	-0.250	-0.007	0.006
or more	(0.090)	(0.391)	(0.019)	(0.053)	(0.391)	(0.014)	(0.009)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.220	0.683	0.383	0.193	0.457	0.473	0.170

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table C.4: Alternative pollution measure: exposure to 3 days of pollution with different aerosol index thresholds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age z-score	Height	Lung capacity	Cognitive test	Height	Lung capacity	Cognitive test
In-utero, 1-1.5	-0.221*	-0.883	-0.023	-0.056	-0.045	-0.029*	-0.008
	(0.114)	(0.701)	(0.031)	(0.067)	(0.622)	(0.015)	(0.014)
In-utero, 1.5-2.25	-0.631***	-1.662*	-0.029	-0.229	-1.697	0.043	0.002
	(0.193)	(0.893)	(0.040)	(0.168)	(1.381)	(0.049)	(0.035)
In-utero, 2.25	-0.259	-1.835***	-0.097**	0.077	-1.366	0.018	0.028
or more	(0.204)	(0.594)	(0.048)	(0.163)	(0.846)	(0.024)	(0.022)
0-12mo, 1-1.5	-0.117	1.107*	0.017	-0.052	0.241	0.005	-0.013
	(0.119)	(0.583)	(0.022)	(0.075)	(0.492)	(0.013)	(0.019)
0-12mo, 1.5-2.25	-0.533***	0.063	-0.056*	-0.018	1.092	0.045	0.016
	(0.160)	(0.923)	(0.031)	(0.128)	(1.187)	(0.033)	(0.026)
0-12mo, 2.25	-0.401***	-0.255	-0.037	-0.031	-0.894	0.016	-0.019
or more	(0.144)	(0.975)	(0.031)	(0.106)	(0.822)	(0.024)	(0.018)
13-24mo, 1-1.5	-0.076	0.387	-0.020	0.012	-0.129	-0.036**	0.010
	(0.078)	(0.696)	(0.020)	(0.106)	(0.526)	(0.016)	(0.012)
13-24mo, 1.5-2.25	-0.346***	-0.632	-0.061*	0.098	-1.331	0.008	-0.030
	(0.128)	(1.098)	(0.031)	(0.119)	(1.090)	(0.021)	(0.028)
13-24mo, 2.25	-0.242*	0.609	-0.065**	0.141	-0.327	-0.014	0.013
or more	(0.141)	(0.859)	(0.028)	(0.136)	(0.879)	(0.031)	(0.025)
Observations	3,790	4,180	2,386	3,344	3,644	3,627	3,030
R-squared	0.220	0.682	0.382	0.193	0.456	0.473	0.170

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Table C.5: Estimated effects by trimester

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2000		2007			2014	
	Height for age	Height	Lung capacity	Test score	Height	Lung capacity	Test score
Trimester 1	-0.226	-1.138	-0.047	0.047	0.047	-1.666	0.044
	(0.294)	(1.022)	(0.049)	(0.153)	(0.153)	(1.559)	(0.043)
Trimester 2	-0.167	-1.368	-0.018	-0.015	-0.015	-0.034	0.000
	(0.224)	(1.403)	(0.059)	(0.101)	(0.101)	(1.344)	(0.034)
Trimester 3	-0.413*	-1.567	-0.055	-0.184	-0.184	-1.647	0.029
	(0.212)	(1.314)	(0.037)	(0.211)	(0.211)	(1.427)	(0.036)
0-12mo	-0.378***	-0.697	-0.043*	0.003	0.003	-0.186	0.030
	(0.106)	(0.691)	(0.023)	(0.083)	(0.083)	(0.947)	(0.026)
13-24mo	-0.247***	-0.140	-0.054***	0.114	0.114	-1.065	-0.004
	(0.089)	(0.710)	(0.020)	(0.098)	(0.098)	(1.026)	(0.028)
Observations	3,790	4,180	2,386	3,344	3,344	3,644	3,627
R-squared	0.218	0.681	0.359	0.192	0.192	0.179	0.158

Notes: All analyses include district, children's year and month of birth fixed effects. Standard errors are clustered at the grid level (pollution exposure area). * p<0.1, ** p<0.05, *** p<0.01.

Selective mortality and the long-term effects of early-life exposure to natural disasters*

Margaret Triyana † and Xing Xia ‡

July 31, 2018

Abstract

We analyze the effects of early-life shocks with varying degrees of severity on mortality and human capital outcomes in the Philippines. We exploit variations in typhoon exposure and sharp increases in short-term disaster relief efforts in the 1960s. Severe typhoons are associated with increased mortality and adverse long-term outcomes. Before the increase in disaster relief efforts, mortality from in utero exposure to severe typhoons was 10 percent; survivors exhibited similar levels of human capital as the unaffected. After the increase in disaster relief, the mortality effects were mitigated; however, survivors exhibited lower human capital in the long term.

Keywords: fetal origins hypothesis, selective mortality, long-term outcomes, Philippines, natural disasters, disaster relief

JEL codes: I12, I15, O15

*We are grateful to Julien Labonne, Douglas Almond, Bhashkar Mazumder, John Ham, Jessica Pan, Ethan Lieber, Joe Kaboski, seminar participants at Yale-NUS, NUS, University of Notre Dame, participants in the 2017 Conference on Health and Development and the 2018 CeMENT Mentoring Workshop for comments and suggestions. Michelle Quak provided invaluable assistance with GIS. Rais Kamis, Jason Carlo Carranceja, and Adila Sayyed provided excellent research assistance. This research was financially supported by a Yale-NUS College Start-up Grant, a NUS HSS Seed Fund, and a Nanyang Technological University Start-up Grant.

†Keough School of Global Affairs, University of Notre Dame, 3169 Jenkins Nanovic Hall, Notre Dame IN 46556. Email: mtriyana@uchicago.edu

‡Yale-NUS College, 28 College Avenue West 01-501, Singapore 138533. Email: xing.xia@yale-nus.edu.sg

1 Introduction

Environmental shocks often have lasting consequences on human capital development. Early-life exposure to negative environmental shocks may increase mortality (Doocy et al., 2007; Jayachandran, 2009) and have negative effects on human capital that persist into adulthood (Almond and Currie, 2011; Currie and Vogl, 2013). Additionally, climate change is expected to intensify the frequency and unpredictability of environmental shocks. Much of the existing literature focuses on the effects of rare and severe natural disasters. However, the question remains: with repeated environmental shocks, how do the effects vary by severity and frequency?

In this paper, we examine the effects of early-life exposure to environmental shocks of varying intensity on human capital outcomes in a developing country, where geography and poverty present unique challenges. Environmental shocks occur more frequently in developing than in high-income countries (Hsiang and Jina, 2014), and the adverse effects of such shocks may be more pronounced in developing countries (Currie and Vogl, 2013; Hanna and Oliva, 2016; Arceo et al., 2016). Moreover, studies on developing countries are often plagued by the paradox of mortality selection: if the strong are more likely to survive, higher mortality rates could correspond to better adult outcomes among the positively selected survivors (Deaton, 2007; Bozzoli, Deaton, and Quintana-Domeque, 2009). Hence, if an early-life shock leads to higher mortality, its long-term effects on human capital may not be observed (Currie and Vogl, 2013; Meng and Qian, 2009). We analyze changes in both mortality and observed long-term outcomes after a sharp increase in post-disaster relief efforts, thereby illustrating the challenge in evaluating the effectiveness of post-disaster aid due to the paradox of mortality selection.

Specifically, we use the unpredictability of severe tropical cyclones – typhoons – in the Philippines as a natural experiment to examine the persistent effects of early-life typhoon exposure. The Philippines is prone to typhoons: on average, five to six typhoons make

landfall every year, one of which is a severe typhoon (Saffir-Simpson scale 4 or 5).[1] Despite the high frequency of natural disasters in the Philippines, disaster relief funds were virtually non-existent before Ferdinand Marcos came to power in 1965. The Marcos regime introduced various measures to increase funding for short-term disaster relief, which aimed to protect the affected population from some of the immediate deleterious effects of natural disasters. If the relief efforts implemented during Marcos' regime reduced typhoon-induced early-life mortality, we may observe changes in the long-term scarring among survivors. We find evidence consistent with the hypothesis that the greater post-disaster relief efforts protected exposed children from mortality, and the lower mortality, in turn, changed the survivors' observed long-term human capital outcomes.

Our empirical strategy exploits variations in the timing, geographic path, and intensity of typhoons in the Philippines. The high degree of similarity among different typhoon occurrences allows us to estimate the dose-response effects of early-life shocks by comparing the effects of a severe (Saffir-Simpson scale 4 or 5) typhoon to those of a less intense (Saffir-Simpson scale 1, 2, or 3) one. We employ a difference-in-differences method to analyze the effects of in utero and post-natal exposure to severe and less intense typhoons. We then separate early-life typhoon exposure into two periods: before and after Marcos assumed office in December 1965 to examine the increase in post-disaster relief efforts. Crucially, the 1990 Philippine Census recorded individuals' year and municipality of birth, which allows us to match each individual with the typhoon activities in his/her municipality of birth while he/she was in utero and in his/her first two years of life. We combine data from the 1990 Census and historical data on the intensity and paths of all tropical storms passing through the Philippines from 1945 to 1990. To analyze the effects on mortality, we use cohort size by birth place and the fraction of males in each cohort as proxies for mortality. To study the effects on long-term human capital accumulation, we use educational attainment and occupation in 1990 as outcome variables.

[1] Author's calculations based on tropical cyclone data from 1945 to 2015.

We find strong dose-response effects: large negative effects from severe typhoons, but not from less intense typhoons. In utero exposure to severe typhoons is associated with a 7 percentage point reduction in male cohort size (by birth place) and a lower fraction of males. These results are consistent with the fragile male hypothesis, which argues that males are more vulnerable to negative early-life shocks than females (Kraemer, 2000; Mathews et al., 2005). The large reduction in male cohort size, but small effects on female cohort size confirms that the reduction in cohort size is driven by an increase in early-life mortality rather than a decline in fertility. Less intense typhoons had little effect on either cohort size or the fraction of males. We then analyze the effects of early life exposure to typhoons on survivors' educational and labor market outcomes and find that only severe typhoons are associated with substantial adverse outcomes.

We also find changes in both mortality and the long-term effects on survivors after the Marcos regime came to power. The Marcos regime introduced various disaster relief efforts, most notably the establishment of a coordinating body to administer short-term disaster relief efforts. Before the Marcos regime, in utero exposure to severe typhoons reduced cohort size by almost 10 percentage points but had little impact on long-term outcomes – the surviving cohorts on average exhibited similar human capital outcomes as the unaffected cohorts. In contrast, under the Marcos regime, when disaster relief funds became more available, albeit still limited, in utero exposure to severe typhoons had little impact on mortality, but significantly reduced survivors' educational attainment and occupational skill level – 0.3 fewer years of education (with larger effects on males) and a 15 percentage point lower probability of attaining a skilled occupation. These findings suggest that the post-disaster relief efforts provided the social protection necessary to mitigate the adverse effects on early-life mortality, but were not sufficient to alleviate the long-term scarring effects. Under the Marcos regime, some individuals who would have died survived and adverse long-term effects are observed in these individuals. In other words, deteriorating long-term outcomes after the increase in disaster relief efforts actually reflect improved chances of

survival.

We compare siblings within the same household to explore the role of post-natal parental investments. Using household fixed effects, the adverse long-term effects are intensified. The estimates with household fixed effects are larger in magnitude than the difference-in-differences estimates with municipality fixed effects. This suggests that post-natal parental investments may reinforce the differences between siblings caused by the negative pre-natal shocks, which is consistent with the findings in Almond et al. (2009). We also examine heterogeneity in the long-term effects by family socioeconomic status (SES). The adverse effects are more pronounced on low-SES families than on high-SES families, suggesting that family resources affect their ability to cope with the aftermath of a severe typhoon and engage in compensating behavior for the affected children.

The remainder of the paper is organized as follows. Section 2 discusses the related literature; Section 3 describes typhoons and disaster relief policies under Marcos; Section 4 describes the data; Section 5 presents the estimation strategy and results; Section 6 concludes with policy implications.

2 Literature

Research has shown that negative early-life shocks can affect mortality and have persistent impacts on survivors' later life outcomes (Almond and Currie, 2011; Currie and Vogl, 2013). Children in utero are especially vulnerable because their fetal programming may be altered by negative shocks, leaving them susceptible to poor well-being in adulthood (Barker, 1995).

The competing effects of culling and scarring affect the observed long-term effects among survivors. When the effect of culling dominates, we may observe no scarring because survivors are highly positively selected (Kannisto et al., 1997). Evidence from birth cohort data demonstrates that, in high mortality settings, declines in child mortality are associated with decreased height, but in settings with limited selective mortality, declines in child mortality are associated with increased height, and, in turn, height correlates with skill

(Deaton, 2007; Bozzoli, Deaton, and Quintana-Domeque, 2009). With low selective mortality, scarring among survivors may be more pronounced, such as the case of the 1918 flu pandemic (Almond, 2006; Lin and Liu, 2014). This paper builds on Bozzoli, Deaton, and Quintana-Domeque (2009) and illustrates the competing effects of culling and scarring in a within-country setting where a sharp increase in disaster relief efforts led to changes in mortality selection.

Negative shocks early in life lower children's human capital endowment. This makes later life investments difficult, leading to poor human capital outcomes in adulthood (Cunha and Heckman, 2007). However, it is possible that mediation occurs to mitigate the initial effects of negative shocks. Conditional on survival, early interventions may offer protective effects if they have occurred in the critical period of development (Currie and Almond, 2011; Almond and Mazumder, 2013). They may allow affected children to catch up (Aguilar and Vicarelli, 2011; Adhvaryu et al., 2015; Gunnsteinsson et al., 2016). By analyzing the effects of increased disaster relief efforts, we also extend the literature on the mitigating effects of early-life interventions (Almond et al., 2017).

Our paper is closely related to a body of literature that analyzes the effects of early-life exposure to tropical cyclones – typhoons in the northwestern Pacific and hurricanes in the Atlantic Ocean. We find a strong dose-response effect to early-life exposure to tropical cyclones – severe storms are associated with adverse outcomes, whereas less intense storms are associated with small effects. These findings stand in stark contrast to findings from the U.S. and Brazil, where low-intensity tropical cyclones can have large negative effects on both short and long-term outcomes (Simeonova, 2011; Currie and Rossin-Slater, 2013; Karbownik and Wray, 2016; de Oliveira and Quintana-Domeque, 2016). One plausible explanation for the differences in the findings is the role of adaptation. Due to the high frequency of low-intensity typhoons in the Philippines, residents may be more prepared, both mentally and physically, to cope with low-intensity tropical cyclones. Other evidence from the US also suggests that long-term adaptation mitigates the effects of short-term shocks (Zivin et al.,

2018).

We also add to the literature on the effects of typhoons in the Philippines. Ugaz and Zanolini (2011) find significant reductions in the height of children who were exposed to severe typhoons while in utero. Deuchert and Felfe (2015) and Anttila-Hughes and Hsiang (2013) both suggest that Philippine households experience persistent reductions in household assets and expenditure several years after a typhoon; the reduction in expenditure in turn reduces children's educational attainment (Deuchert and Felfe, 2015) and increases infant mortality rates (Anttila-Hughes and Hsiang, 2013). Compared to these earlier studies, our use of multiple typhoons in a generalized difference-in-differences framework allows us to compare the effects of in utero exposure relative to exposure at other ages. We do not exclude the possibility that typhoon exposure adversely affects children of all ages, perhaps through reductions in household consumption. Rather, we focus on the incremental effects of in utero exposure – the difference between the effects on cohorts who were exposed in utero and cohorts who were exposed at other stages of life. Other than reduced income and expenditure, poor maternal nutrition and sustained maternal stress could also lead to adverse long-term effects (Liu et al., 2016).

3 Context

3.1 Typhoons in the Philippines

The Philippines is the fourth most disaster-prone country in the world.[2] Given its location in the western Pacific Ocean, the Philippines is the largest country to be exposed to typhoons on a regular basis – on average five to six typhoons make landfall in the country annually (Table 1), with an even larger number entering Philippine waters. Typhoons form all year round, the peak months being July to October.

The country comprises about 7,000 islands, which are categorized into the following island

[2]Source: UN, The Human Cost of Weather Related Disasters 1995-2015. URL (last accessed December 27, 2017): http://www.unisdr.org/2015/docs/climatechange/COP21_WeatherDisastersReport_2015_FINAL.pdf

groups to reflect differences in both geographic location and economic development: Northern Luzon, Southern Luzon, Visayas, and Mindanao.[3] These island groups are administratively further divided into provinces and municipalities. The northeastern parts of the country – Northern Luzon, the eastern parts of Southern Luzon, and the Visayas – are especially prone to typhoons.[4] However, most municipalities in the Philippines are prone to natural disasters such as earthquakes and volcanic eruptions. Hence, despite the high prevalence of migration in the Philippines, about 14 percent between 1960-70, it is uncommon for citizens to migrate within the Philippines to avoid natural disasters, including typhoons.

Due to the country's familiarity with typhoons, institutions and infrastructure in the Philippines may be somewhat prepared to cope with low-intensity typhoons, but not severe ones. For example, Franklin and Labonne (2017) show that local labor markets are unaffected by low-intensity typhoons, but substantially disrupted by severe ones. We follow Franklin and Labonne (2017) and categorize typhoons into two intensity groups: low-intensity corresponds to categories 1, 2, and 3 on the Saffir-Simpson scale, and severe corresponds to categories 4 and 5.

3.2 Disaster relief before and under the Marcos regime

Disaster relief gained prominence after Ferdinand Marcos came to power in December 1965. Previously, it had fallen under the National Civil Defense Administration established by the 1954 Civil Defense Act. The body was tasked with protecting the welfare of the population during national emergencies, including wars and natural disasters.[5] However, the planning body was poorly funded and uninterested in disaster relief. In spite of the body's lack of commitment, the President had the executive power to declare a 'state of calamity' to

[3]We divide the Luzon island groups into two parts to separate the predominantly agricultural north from the more industrialized south. They experienced divergent population growth trends during the period we study. Our definition of Northern Luzon includes Ilocos Region, Cordillera Administrative Region, Cagayan Valley, and Central Luzon. Southern Luzon includes NCR (Metro Manila), Southern Tagalog (Calabarzon and Mimaropa), and Bicol Region.

[4]Source: http://vm.observatory.ph/hazard.html (Last accessed December 27, 2017.)

[5]Source: NDCC, History of Disaster Management in the Philippines, URL (last accessed December 27, 2017): http://www2.wpro.who.int/internet/files/eha/tookit_health_cluster/History%20of%20Disaster%20Management%20

provide disaster relief. Filipino presidents before Marcos rarely used their executive power to do so, and to the best of our knowledge, 'state of calamity' proclamations started to include provisions for disaster-relief funds under Marcos.[6] Additionally, apathy towards disaster relief changed after the Casiguran earthquake in August 1968. The 7.6 magnitude earthquake killed at least 207 people, most of whom died in Manila. Following this disaster, Marcos issued Administrative Order No. 151 in December 1968 and the National Committee on Disaster Operation was created.[7] This instigated the coordination of disaster response and disaster relief funding across different agencies. The committee issued a standard operating procedure that prescribed the organizational set-up for disasters from the national level down to the municipal level.

Post-disaster efforts focused on short-term assistance, such as clothing, food, and medicine after the disasters to mitigate immediate adverse outcomes such as disease outbreaks – one potential cause of child mortality in this context. Key stakeholders, including the military, were mobilized to provide disaster relief and to maintain the stability of prices of prime commodities such as rice, the main staple. Marcos also tapped foreign assistance to provide relief goods after major typhoons and personally directed the relief effort, involving his wife and son in some cases (Warren, 2013). There is evidence that these disaster relief efforts were part of the regime's political manipulation to give preferential treatment to their supporters (Warren, 2013). Nonetheless, on average, the availability of post-disaster assistance may mitigate the short-term deleterious effects of negative shocks, such as early-life mortality. To our knowledge, there is no detailed record of how much and where resources were sent after each disaster during the Marcos regime. Therefore, we unfortunately cannot exploit regional variations in post-disaster relief to evaluate the per capita effect of relief spending. Instead, we assess the overall effectiveness of post-disaster relief policies by estimating the

[6]Source: The Official Gazette of the Philippines, last accessed on March 27, 2018.

[7]Post-disaster relief efforts were further strengthened in the next decade. The disaster management plan was augmented after typhoon Sening (Joan), which struck in October 1970 and heavily affected the Metro Manila area. Marcos approved a Disaster and Calamities Plan, which led to the creation of a National Disaster Control Center (NDCC). The NDCC includes almost all Department Secretaries as members and still serves as the most important policy-making agency for disasters in the Philippines.

average changes in typhoon-induced mortality and long-term effects after the introduction of such policies.

4 Data

We draw our outcome variables from the Philippine Census of Population and Housing 1990 (hereafter, CPH 1990). We match each individual in CPH 1990 with historical typhoon exposure information in his or her municipality of birth. To measure typhoon exposure and intensity, we use the best-track data from the Japan Meteorological Agency Tropical Cyclone Database (henceforth, JMA) and the typhoon analogs (TD-9635) collected by the National Climatic Data Center.

4.1 Typhoon Data

The JMA provides the most reliable information of all tropical storms passing the Western North Pacific (WP) basin. For each tropical storm between 1951 and 1990, JMA records the longitude and latitude of the storm center and the minimum central pressure every six hours.[8] The typhoon analogs (TD-9635), collected by the National Climatic Data Center, provide the same information for typhoons that passed through the WP basin between 1945 and 1950.

We use the coordinates of the storm center to identify affected municipalities. First, we generate best-fit lines through the six-hourly observations to identify the storm path. Then, we calculate the distance between the centroid of the municipality and the storm path for each municipality-storm pair. If the distance is within 30 kilometers (km), we treat the municipality as affected by the storm.[9]

To measure storm intensity, we use the minimum central pressure (MCP).[10] The intensity

[8]Links to both data sets can be found on the IBTrACS webpage, URL (last accessed December 27, 2017) https://www.ncdc.noaa.gov/ibtracs/index.php?name=rsmc-data

[9]For robustness, we use a distance of 60 kilometers from the storm center to define affected municipalities. We also use the nearest distance between the municipality and the storm track to measure municipality-to-storm distance. The results are qualitatively similar.

[10]We use the MCP instead of the more commonly used maximum sustained wind speed (MSW) due to

of a storm fluctuates as the storm proceeds, but our databases provide MCP measures only every six hours. To measure the storm intensity for each of the affected municipalities, we use the weighted average of the MCP readings of the two nearest observation points, using the inverse distance between the observation and the municipality as weights. We then categorize typhoon intensity according to the Saffir-Simpson scale[11], and generate an indicator for "small typhoon" if the storm is of category 3 or lower and "severe typhoon" if the storm is of category 4 or 5 when the storm reached the municipality.

Storm path and severity vary considerably across the island groups. For example, Figure A.1 shows the paths and severity of all typhoons that passed the Philippines in 1970. Table 1 shows the number of typhoons per municipality-year by the four island groups. We separate the sample into two periods: pre-1965 and post-1965. On average, the number of small typhoons is similar across the two time periods. However, severe typhoons are more frequent in the post-1965 period, especially in the Luzon island groups.

4.2 Philippine Census Data

Outcome variables Our identification strategy requires us to link individuals' later life outcomes to typhoon exposure in utero and in his or her first two years of life. We use the CPH 1990 10% Household Sample because it is the only census sample that identifies the usual place of residence of each respondent's mother at the time of the respondent's birth. This longer census questionnaire was administered to approximately 10 percent of the population. Our outcomes of interest include cohort size by birth place as well as the education and labor market outcomes of all individuals, including non-resident household

data limitations. MSW was not available in the JMA database until 1972; additionally, the MSW calculation was revised in the 1980s to be consistent across meteorological agencies. When both MSW and MCP are available in the JMA database, the two measures are highly correlated (-0.833, p-value <0.01). Moreover, recent meteorological studies found that due to changes in practice over time at meteorological agencies, records of MSWs for pre-1980s storms in the WP basin are likely to be of low quality (Knapp et al., 2013). Hence, MCP serves as a more accurate measure of storm intensity for pre-1980 tropical cyclones in the WP basin.

[11] A category 5 typhoon is one with MCP below 920 millibars; a category 4 typhoon is one with MCP between 920 and 944 millibars; category 3 is between 945 and 964 millibars; category 2 is between 965 and 979 millibars; category 1 is between 980 and 999 millibars. Storms with MCP at or above 1000 millibars are not considered typhoons.

members such as overseas workers. We draw these variables from the census.

Ideally, detailed data on fetal, infant, child, and adult mortality rates would be available for each cohort born in a given municipality. Unfortunately, such mortality records, especially fetal mortality, do not exist for the Philippines. We follow Jayachandran (2009) and infer mortality by measuring cohort size in 1990 based on individuals' municipality of birth. As such, cohort size by birth place is proxied by the probability of survival to May 1, 1990.[12]

Educational outcomes include literacy, high school completion, and years of education. The census does not include information on respondents' labor market earnings but provides detailed information on respondents' occupations. Based on each respondent's reported occupation, we construct three indicators of occupational skill level: whether the respondent has a skilled occupation, an associate professional occupation, or a professional occupation. The Data Appendix details the construction of cohort size by birth place, years of education, and occupational skill level indicators.

Exposure variables To identify whether an individual was affected by typhoons in early-life, we link each respondent in CPH 1990 to the typhoon data according to the respondent's age and municipality of birth. Municipality of birth is given by the usual place of residence of the respondent's mother at the time of the respondent's birth.

Our main exposure variables are the expected number of small and severe typhoons that each respondent (or cohort) was exposed to during the in utero period and in the first two years of life.[13] We use this expected number of typhoons as our treatment variable because respondents' month of birth is unfortunately not available. However, we do observe each

[12]The reductions in observed cohort size is due to a combination of early-life (under the age of one) and later-life mortality. While the highest mortality is attributed to early-life deaths, later-life mortality may be a concern. To address this, we replicate our cohort size analysis using cohort size in the 1970 Census (CPH 1970). CPH 1970 allows us to restrict the sample to cohorts born before 1965 (aged between 5 and 23 in 1970), but CPH 1970 provides only the respondent's province of birth, not the municipality of birth. Hence, typhoon exposure is defined at the province level, which severely limits the accuracy of the estimates. Nonetheless, the results using CPH 1970 are similar in magnitude to our main results using CPH 1990.

[13]About 95% of siblings were born in the same municipality, so migration in early-life is low. Therefore, we expect measurement error in our exposure variable in the first two years of life to be small. We explore this further in the Appendix.

respondent's age as of May 1, 1990, as well as the exact date that the typhoons passed his or her municipality of birth. To construct the expected number of typhoon exposures, we assume that an individual of age y is equally likely to be born any day between May 2, $1990 - y - 1$, and May 1, $1990 - y$, and that gestation starts 40 weeks prior to the potential date of birth.[14] We then use the date that a typhoon passed the municipality of birth to construct the probability that the individual was exposed to the typhoon in utero. We also construct the probability of exposure in the first two years of life in a similar fashion. We sum the probabilities across typhoons to derive the expected number of typhoons that each respondent (or cohort) is exposed to at each stage of life. The Data Appendix details the construction of the exposure variables.

This measure allows us to fully exploit the temporal variations of typhoons. For example, because respondents' ages are recorded as of May 1, 1990, a typhoon that took place between August and October 1970 could potentially affect the in utero period of two cohorts – ages 18 and 19 – with a higher probability of affecting the 19-year-old cohort in utero than the 18-year-old cohort. In contrast, a typhoon that took place in May or June of 1970 could only possibly affect the in utero period of one age cohort – the 19-year-old cohort. These variations are fully captured by our exposure variable: expected number of typhoons, but would not be captured by a dummy variable indicating whether a typhoon took place one year before birth or the year of birth. The temporal dimension is especially important since most typhoons in the Philippines take place between June and November of each year.[15]

Sample restrictions Respondents in the Philippines should finish high school by the age of 16 and vocational college by the age of 18 if there is no grade repetition or late enrollment. We restrict the sample to respondents over the age of 18 to ensure that most would have

[14] Using 36 and 38 weeks of gestation as alternative specifications yields qualitatively similar results.

[15] We perform all our analysis with two sets of alternative typhoon exposure variables: (1) dummy variables indicating whether *any* small or severe typhoon passed the municipality of birth in the three years around birth, and (2) the *number* of small or severe typhoons that passed the municipality of birth in the three years around birth. Results using these alternative measures are consistent with our main findings (Section 5.5).

entered the labor market by 1990.[16] As we are interested in typhoon exposure in the first 3 years after conception, we further restrict the sample to respondents between the ages of 18 and 43 (those born between 1947 and 1972) so that we have each respondent's full history of typhoon exposure in utero and in his first two years of life.[17]

5 Results

5.1 Estimation strategy

We exploit the temporal and geographical variation of the typhoon paths as well as exogenous variations in typhoon intensity. We employ a difference-in-differences method, exploiting variations at the cohort-municipality level. Specifically, we compare individuals who were exposed to typhoons either in utero or in their first two years of life to those who were either born in the same municipality in a different year or born in a different municipality in the same year and were, therefore, exposed neither in utero nor in the first two years of their lives.

Subsection 5.2 presents the results on mortality, using cohort size by birth place and the fraction of males in each cohort as outcome variables to infer mortality. Subsection 5.3 presents the results on long-term educational and occupational outcomes. In subsection 5.4, we extend our analysis to sibling comparisons. We compare individuals who were exposed to typhoons either in utero or in their first two years of life to a sibling who was born in a

[16]The youngest cohorts may still be in tertiary education in 1990. To assess the extent of this concern, we use individuals' completed years of schooling in the 2000 Census of Population and Housing to infer the percentage of each cohort that would have completed schooling by 1990 if there had been no gaps or repetitions in their schooling. The data indicate that 86 percent of the 18-year-old cohort, 89 percent of the 19-year-old cohort, 94 percent of the 20 and 21-year-old cohorts, and 95 percent of the 22-year-old cohort would have completed schooling by 1990.

[17]Table A.1 shows the outcomes of interest for those who were exposed to typhoons in utero either before (Panel A) or after 1965 (Panel B). We divide the sample to those exposed in utero to small typhoons only (column 1), those ever exposed to severe typhoons (column 2), and those never exposed to any typhoons (column 3). Conditional on survival to adulthood, the educational and occupational outcomes for those who were exposed to severe typhoons *before 1965* are *higher* than those who were exposed to small typhoons or never exposed to typhoons (Table A.1), but those who were exposed to severe typhoons *after 1965* have lower occupational status.

different year and, hence, never exposed during the early-life period.

5.2 Cohort size and mortality

We begin by analyzing the effects of early-life exposure to typhoons on mortality using cohort size by birth place and the fraction of males as outcome variables. The estimated effects capture the cumulative effects of typhoon exposure on fetal, infant, child, and adult mortality. We estimate the following equation for each birth-municipality m and birth-year t,

$$\begin{aligned}\ln(CohortSize_{mt}) = {} & \theta_0\ small\_inutero_{mt} + \theta_1\ small\_postnatal_{mt} \\ & + \beta_0\ severe\_inutero_{mt} + \beta_1\ severe\_postnatal_{mt} \\ & + \phi_m + \tau_t \times \psi_{island} + \gamma_{region} \times t + \epsilon_{mt} \end{aligned} \tag{1}$$

where $CohortSize_{mt}$ refers to the number of individuals *born* in municipality m and year t who survived until May 1, 1990. The treatment variable $small\_inutero$ measures exposure to small typhoons in utero, and $small\_postnatal$ measures exposure to small typhoons in the first two years of life; $severe\_inutero$ and $severe\_postnatal$ are the corresponding measures for severe typhoons.

In all subsequent analysis, we include municipality fixed effects, ϕ_m, to take into account non-time varying municipality characteristics. We also include birth-year by island group fixed effects, $\tau_t \times \psi_{island}$, to account for differences in education and economic development policies across the four island groups that may affect the outcome of interest. In addition, we include region-specific time trends, $\gamma_{region} \times t$, to allow for differential population growth trends in the thirteen regions.[18] Standard errors are clustered two-ways at both the municipality and the province-by-birth-year levels. Clustering by province-birth-year allows us to

[18]There are 13 regions in the Philippines across the 4 island groups in 1990: 3 in the Northern Luzon island group (Ilocos Region, Cagayan Valley, and Central Luzon), 3 in Southern Luzon (National Capital Region, Southern Tagalog, and Bicol), 3 in Visayas (Western Visayas, Central Visayas, and Eastern Visayas), and 4 in Mindanao (Western Mindanao, Northern Mindanao, Southern Mindanao, and Central Mindanao).

account for the spatial correlation across municipalities in the typhoon exposure measures.

One caveat of using cohort size to study mortality effects is that cohort size also reflects changes in conception rate. To address this concern, we also use the fraction of males in each cohort and each municipality of birth as an outcome variable. Under the fragile male hypothesis (Kraemer, 2000), males are more susceptible to fetal and infant mortality than females. Adverse early-life shocks would reduce the fraction of males. If, instead, the adverse shock reduces the rate of conception, we would expect similar reductions in the male and female cohort size and no change in the fraction of males. Therefore, reductions in the fraction of males would provide suggestive evidence that the changes in cohort size are likely to be attributable to changes in mortality, rather than changes in the conception rate.

Table 2 presents the results of estimating Equation 1. We restrict our sample to cohorts between ages 2 and 43 in columns 1 to 4.[19] In columns 5 to 8, we further restrict our sample to cohorts aged 18 to 43 to match the sample we use for education and occupational outcomes. We find that severe typhoons are associated with increased mortality, but small typhoons are not. On average, in utero exposure to a severe typhoon reduces cohort size (by birth place) by about 5 percent, but exposure to a small typhoon has no significant effect on cohort size (column 1). The estimated effect is similar, about 7 percent, when we restrict the sample to ages 18 to 43 (column 5). We find no significant effects on cohort size resulting from exposure to small or severe typhoons in the first two years of life. Our finding is consistent with the literature on the adverse effects of in utero exposure to severe negative shocks and suggests a strong dose-response effect to the severity of the shock.

In addition, consistent with the fragile male hypothesis (Kraemer, 2000) and earlier empirical evidence (Mathews et al., 2005; Almond and Mazumder, 2011), the mortality effects of exposure to severe typhoons are more pronounced among males – the effect on male cohort size is large (8 percent) and statistically significant (column 2), whereas the effect on female cohort size is much smaller (3 percent) and statistically insignificant (column 3).

[19]We exclude cohorts aged below 2 or above 43 because we do not have information on typhoon exposure in early life for those cohorts.

The effects on the fraction of males (columns 4 and 8), albeit imprecisely estimated, indicate that exposure to a severe typhoon reduces the fraction of males in each cohort by 1 to 2 percentage points, suggesting that the reduction in male cohort size is driven by increases in early-life mortality rather than reductions in the conception rate.

We then examine the changes in mortality under the Marcos regime by separating typhoon exposures that occurred before and after Marcos assumed office in December 1965 and estimate the following equation:

$$\begin{aligned}\ln(CohortSize_{mt}) = {} & \rho_0\ pre\_65\_small\_inutero_{mt} + \rho_1\ pre\_65\_small\_postnatal_{mt} \\ & + \alpha_0\ pre\_65\_severe\_inutero_{mt} + \alpha_1\ pre\_65\_severe\_postnatal_{mt} \\ & + \theta_0\ post\_65\_small\_inutero_{mt} + \theta_1\ post\_65\_small\_postnatal_{mt} \\ & + \beta_0\ post\_65\_severe\_inutero_{mt} + \beta_1\ post\_65\_severe\_postnatal_{mt} \\ & + \phi_m + \tau_t \times \psi_{island} + \gamma_{region} \times t + \epsilon_{mt} \qquad (2)\end{aligned}$$

where the treatment variables $small\_inutero$, $small\_postnatal$, $severe\_inutero$, and $severe\_postnatal$ are interacted with either a pre_65 or a $post\_65$ dummy variable indicating whether the typhoon exposure took place before or after December 1965. The implicit assumption is that all typhoons after December 1965 are covered by Marcos' disaster relief policies. We present the results using December 1968, the month that the National Committee on Disaster Operation was established, as the alternative cut-off in Subsection 5.5.

Table 3 presents the results of estimating Equation 2. The results show stark contrasts of mortality effects before and after 1965. In utero exposure to a severe *pre-1965* typhoon is associated with a statistically significant 10 percent decrease in overall cohort size (by birth place) in 1990 (column 1). This effect is stronger among males (14 percent) than among females (6 percent and not statistically significant). In contrast, the effects of in utero exposure to a severe *post-1965* typhoon are both substantively and statistically insignificant (columns 1-4). The estimates are similar when we restrict the sample size to cohorts between

the ages of 18 and 43 (columns 5 to 8).[20] Although the effects on the fraction of males are imprecisely estimated, the difference in magnitude is large between pre- and post- 1965 severe typhoons – in utero exposure to a severe *pre-1965* typhoon reduces the fraction of males by 2.2 percentage points, whereas a severe *post-1965* typhoon reduces the fraction of males by 0.6 percentage points (columns 4 and 8). On the other hand, the effects of in utero exposure to a small typhoon are insignificant in magnitude both before and after 1965. Exposure to small typhoons in one's first two years of life is associated with a small reduction in cohort size post-1965.[21] The effects of exposure to severe typhoons in one's *first two years of life* are neither statistically nor substantially significant in the two time periods.

We interpret these as suggestive evidence that the Marcos regime's post-disaster relief measures reduced the short-term adverse effects of typhoons. This change in early-life mortality could subsequently affect the observed long-term effects on the survivors, which we turn to next.

5.3 Educational attainment and occupational skill level

Conditional on survival, we estimate the long-term effects of exposure to small and severe typhoons. We estimate the following equation for individual i, born in municipality m and year t:

$$\begin{aligned} y_{imt} = {} & \rho_0 \, pre\_65\_small\_inutero_{mt} + \rho_1 \, pre\_65\_small\_postnatal_{mt} \\ & + \alpha_0 \, pre\_65\_severe\_inutero_{mt} + \alpha_1 \, pre\_65\_severe\_postnatal_{mt} \\ & + \theta_0 \, post\_65\_small\_inutero_{mt} + \theta_1 \, post\_65\_small\_postnatal_{mt} \\ & + \beta_0 \, post\_65\_severe\_inutero_{mt} + \beta_1 \, post\_65\_severe\_postnatal_{mt} \\ & + \delta X_{imt} + \phi_m + \tau_t \times \psi_{island} + \gamma_{region} \times t + \epsilon_{imt} \end{aligned} \tag{3}$$

[20] We restrict the sample to municipalities with at least one male and one female in each age cohort under 43 to ensure a balanced panel for all four outcome variables. Appendix Table A.3 presents results including all municipalities. The results are qualitatively similar.

[21] We explore this further by analyzing the effects of each typhoon category (Table A.15).

where y_{imt} is the outcome of interest. We include the same set of typhoon exposure variables as in Equation 2, with the implicit assumption that once a municipality is exposed to a typhoon, everyone residing in the municipality is exposed. Additionally, we include a male indicator as a covariate (X_{imt}). When occupation is the outcome variable, we also include years of education as a covariate in some specifications. As in Equation 2, we include municipality fixed effects, ϕ_{muni}, birth-year-by-island-group fixed effects, $\tau_t \times \psi_{island}$, and region-specific time trends, $\gamma_{region} \times t$. Standard errors are clustered two-way at both the municipality and the province-by-birth-year levels. In addition, observations are weighted by the person weights provided in CPH 1990.

Education Table 4 presents the effects of early-life exposure to typhoons on educational attainment. There is a strong dose-response relationship with the intensity of the typhoon as well as a sharp increase in the scarring effects after 1965. Exposure to *small* typhoons, both pre- and post-1965, had little effect on educational attainment.[22] In contrast, in utero exposure to *severe* typhoons is associated with adverse effects on all three outcomes. Moreover, the adverse effects of severe typhoons are substantially more pronounced post-1965 than pre-1965. In utero exposure to severe pre-1965 typhoons had little effect on the educational attainment of survivors, which is not surprising since pre-1965 typhoons were associated with high mortality rates. In contrast, in utero exposure to a severe post-1965 typhoon reduces the probability of being literate by 0.771 percentage points (a 0.80 percent reduction from the mean), reduces completed years of education by 0.342 years (a 3.53 percent reduction from the mean), and reduces the probability of completing high school by 1.81 percentage

[22]The estimated effects of in utero exposure to *small,* pre-1965 typhoons are positive for all three outcome variables. Although it seems odd to find long-term positive effects associated with exposure to small typhoons, the magnitudes of the coefficients are small compared to the effects of severe typhoons. In addition, small, non-destructive typhoons may benefit agriculture and fishing by increasing rainfall, lowering temperatures, and raising the abundance and variety of fish in nearshore waters (Lam et al., 2012; Yu et al., 2013). If some small typhoons increase agricultural and fishing income without much damage to local infrastructure, they may improve mothers' and children's nutritional intake in the ensuing months and contribute to positive long-term human capital outcomes. Appendix Table A.16 confirm that these positive effects stem from the lowest category (SS scale 1) typhoons, which are relatively common and non-destructive in the Philippines.

points (a 3.49 percent reduction from the mean).[23] We find no significant effects on education among those exposed to typhoons in their first or second year of life.

Occupation Table 5 presents the effects on occupational skill level. Again, we observe a strong dose-response relationship by the intensity of the typhoon and a sharp increase in the scarring effects after 1965. Exposure to small typhoons, both pre- and post-1965, is associated with small positive gains in occupational skill level.[24] In contrast, in utero exposure to a severe *post-1965* typhoon reduces occupational skill level – it reduces the probability of attaining a skilled occupation by 15.9 percentage points, an associate professional occupation by 6.9 percentage points, and a professional occupation by 5.19 percentage points. We find that exposure to severe *post-1965* typhoons in the first two years of life is also associated with lower occupational skill level, although the magnitude is smaller than in utero exposure.

The effects on labor market outcomes remain when we condition on years of completed education, which is similar to the US case (Karbownik and Wray, 2016). These findings suggest that scarring in the labor market operates through channels other than education, plausibly because educational attainment relies more on cognitive ability whereas labor market productivity relies on both cognitive ability and health. Interestingly, in utero exposure to severe, *pre-1965* typhoons is associated with a small, but statistically significant increase in the probability of attaining an associate professional and professional occupation, suggesting that survivors are extremely positively selected in the pre-Marcos era.

Heterogeneity by gender We explore the heterogeneous effects by gender in Tables 6 and 7 to further explore the fragile male hypothesis (Kraemer, 2000). We find that exposure

[23] Appendix Table A.4 shows the corresponding effects on education pooling all cohorts together. Exposure to small typhoons has no significant effect on educational outcomes, but exposure to severe typhoons is associated with adverse later life outcomes. On average, in utero exposure to one severe typhoon lowers educational attainment by 0.178 years and such exposure in the first two years of life reduces educational attainment by 0.059 years.

[24] Please refer to footnote 22 for an explanation of why small typhoons may have small positive effects. Exposure to severe typhoons is associated with lower occupational skill when we pool the cohorts (Table A.5).

to severe typhoons has a larger effect on male mortality rates than on female mortality rates in both the pre- and post-1965 periods. If surviving individuals are positively selected, the fragile male hypothesis predicts that the long-term scarring effects are larger on males when mortality rates are low; however, we may see no differential long-term effects by gender, or even larger long-term effects on females when early-life mortality rates are much higher among males than among females.

Table 6 presents the heterogeneous effects on educational attainment.[25] Severe, pre-1965 typhoons had no statistically significant effect on literacy or years of education for either gender; males who were exposed to severe typhoons in utero were less likely to complete high school. Post-1965, when the effects on mortality were lower, the effects of in utero exposure to severe typhoons are larger on males than on females for all three educational outcomes (although not statistically significant for the probability of high school completion). These results are consistent with the fragile male hypothesis – when mortality rates are low, long-term scarring is more pronounced among males. Table 7 presents the heterogeneous effects on occupational skill levels. Before 1965, the effect of in utero exposure to severe typhoons on occupational skill levels is not statistically significant for males and slightly positive for females. Post-1965, the adverse effect of in utero exposure to severe typhoons is similar for males and females, with larger effects on the probability that males obtain a skilled occupation and on the probability that females obtain an associate professional or professional occupation.

5.4 Sibling comparisons

We compare those who are exposed to typhoons to their siblings by restricting our sample to households with adult co-resident children. We extend the difference-in-differences framework to this sample using household fixed effects instead of municipality fixed effects. Within-household analysis adds to our study in three ways.

[25]Tables A.6 and A.7 present the results without separating the pre and post-1965 exposure. Severe typhoons are associated with lower education and occupational skill level.

First, this approach controls for unobserved heterogeneity across households. If the unobserved characteristics of households residing in typhoon-exposed areas deteriorate over time, perhaps due to migration, then our identifying assumption would be violated and we may over-estimate the effects of typhoons, especially the post-1965 typhoons. Sibling comparisons address these concerns by controlling for heterogeneities across households.

Second, by comparing the results with household fixed-effects to the results with municipality fixed-effects, we can provide some evidence of post-natal parental investment behavior (Almond et al., 2009). If within-household sibling comparisons yield smaller effects than cross-household difference-in-differences analyses, parents may have compensated for the negative pre-natal shocks by investing more in the affected child after birth (or there may be changes over time in the unobserved heterogeneity of typhoon-exposed households). However, if within-household sibling comparisons yield larger effects than the cross-household analyses, it suggests that parents reinforce negative pre-natal shocks by investing less in the affected child post-natally.

Third, we further stratify our household sample by parental socioeconomic status to examine heterogeneities in the effects of typhoons since low-income households may be more vulnerable to typhoons. In the Philippines, low-income households are more likely to have make-shift houses which may be heavily damaged or destroyed by severe typhoons. Wealthy households are more likely to live in concrete buildings on higher ground, which may be less damaged or remain intact after a severe typhoon. Past research shows that typhoons are more damaging to low-income households' assets than those of high-income households (Anttila-Hughes and Hsiang, 2013; Deuchert and Felfe, 2015; Huigen and Jens, 2006), and wealthy, politically connected families may also be better able to obtain post-disaster funds (Atkinson et al., 2014). We, therefore, expect in utero typhoon exposure to be more damaging to children born to low-income families.

We now restrict our sample to households with adult co-resident children – individuals between the ages of 18 and 43 who reside in the same household as their parent(s) and whose

reported relationship to the household head is either that of "son" or "daughter". We further restrict the sample to individuals with at least one other sibling living in the same household. These restrictions allow us to identify siblings and their parents, but also leave us with a selected sample.[26]

Tables 8 and 9 present the effects of early-life typhoon exposure on education and occupation, respectively, using the adult co-resident children sample. In both tables, columns 1 to 3 present sibling comparison results using household fixed effects. Because this co-resident sample is different from the full sample used in the main analysis, we repeat the difference-in-differences analysis with municipality fixed effects on this co-resident sample (columns 4 to 6) for ease of comparison. The basic patterns found in the cross section persist when we use household fixed effects. Moreover, comparing columns 1 to 3 to columns 4 to 6 in both tables shows that the negative effects of in utero exposure to severe typhoons are larger in magnitude when using household fixed effects than when using municipality fixed effects. This is true for both pre- and post-1965 typhoons. These reduced-form estimates offer suggestive evidence that post-natal parental investment may reinforce the differences between siblings caused by negative pre-natal shocks, but we cannot ascertain the channels through which parents invest differentially in children.

Next, we stratify our sample by household socioeconomic status. Ideally, we would use a measure of family income or wealth around the time of the child's birth as the yardstick to divide households into low and high-SES sub-samples. In the absence of a direct measure of past household income or wealth, we divide the sample according to whether the household head has a skilled occupation.[27] Tables 10 and 11 present the effects of early-life exposure

[26]Respondents in this adult co-resident children sample are on average more educated than the overall population. Average education in the main analyzed sample is 9.4 years, compared to 10.5 years in the adult co-resident household sample. In addition, the fraction of the co-resident sibling sample who were ever exposed pre-1965 are higher than the fraction of the main sample. The post-1965 co-resident sibling sample is more comparable to the post-1965 sample in our main analysis.

[27]We choose this measure because the distribution of the household heads' occupation is relatively stable across age groups, whereas other variables such as years of education vary more by age group. Since household heads in the adult co-resident children sample are older (with an average of 54 years versus 44 years in the CPH1990), the household head's occupation provides a time-consistent way of defining household SES.

to typhoons on education and occupational skill level by household SES. For both years of education and occupational status, the adverse effects are lower among affected children from high-SES households than those from low-SES households.

These reduced form estimates are consistent with two potential mechanisms that lie beyond the scope of this paper. One possibility is that high-SES families have more resources to cope with severe typhoons. Hence, the wealthier the family, the lower a typhoon's impact on the mother's psychological well-being and nutritional intake. Another possibility is that parents in high-SES families engage in post-natal compensating behavior – enhancing human capital investment in a child who experienced negative shocks in early life, whereas parents in low-SES families with limited resources adopt reinforcing behavior post-natally – reducing investment in a child who experienced negative shocks in early life.

5.5 Robustness

Effects by timing of exposure We conduct an event study analysis to show the effects of exposure to typhoons that took place before, during, and after the gestation period. To keep the model parsimonious, we include only severe typhoons in this analysis. We estimate the following equation:

$$\begin{aligned}\ln(cohort\ size_{mt}) = {} & \alpha_{-1}\ pre\_65\_severe_{m,t-3\ or\ t-2} + \alpha_0\ pre\_65\_severe_{m,t-1\ or\ t} \\ & + \alpha_1\ pre\_65\_severe_{m,t+1\ or\ t+2} + \alpha_2\ pre\_65\_severe_{m,t+3\ or\ t+4} \\ & + \beta_{-1}\ post\_65\_severe_{m,t-3\ or\ t-2} + \beta_0\ post\_65\_severe_{m,t-1\ or\ t} \\ & + \beta_1\ post\_65\_severe_{m,t+1\ or\ t+2} + \beta_2\ post\_65\_severe_{m,t+3\ or\ t+4} \\ & + \phi_m + \tau_t \times \psi_{island} + \gamma_{region} \times t + \epsilon_{mt} \qquad (4)\end{aligned}$$

where $pre\_65\_severe_{m,t-3\ or\ t-2}$ is a dummy variable indicating whether the birth-municipality, m, was exposed to any pre-1965 severe typhoons two to three years before the individual's

birth-year, t; $pre\_65\_severe_{m,t-1\ or\ t}$ indicates whether the birth-municipality, m, was exposed to any pre-1965 severe typhoons either one year before or during the birth-year, t; and similarly for the other treatment variables. We use two-year windows here to avoid collinearity and to reduce the number of coefficients we have to estimate. By construction, α_0 and β_0 capture the effects of in utero exposure to severe typhoons, since exposure that took place one year before or during the birth year is possibly in utero exposure. α_{-1} and β_{-1} measure the effects of severe typhoons that took place before conception. α_1, α_2, β_1 and β_2 reflect the effects of post-natal exposure to severe typhoons.

Figure 1 shows the results of estimating Equation 4. Each panel of Figure 1 shows the results of separate regressions where the outcome variables are the fraction of males for Panel A, ln(*cohort size*) for Panel B, ln(*male cohort size*) for Panel C, and ln($female\ cohort\ size$) for Panel D. In each panel, we plot the coefficient estimates and 95% confidence intervals of the key coefficients of interest from Equation 4. Coefficients of the pre-1965 typhoons: α_{-1}, α_0, α_1, and α_2, are plotted on the left-hand side of each panel; coefficients of the post-1965 typhoons: β_{-1},β_0, β_1, and β_2, are plotted on the right-hand side of each panel. Estimates for these regressions are also presented in Appendix Table A.8.

The contrast between pre-1965 and post-1965 typhoon exposures is evident in Figure 1. The estimates for α_0 are negative and statistically significant for three outcome variables: the fraction of males, cohort size, and male cohort size. In contrast, the estimates for β_0 are almost zero and statistically insignificant for all four outcome variables. These findings are consistent with our previous results that severe, pre-1965 typhoons substantially reduced cohort size, especially male cohort size, whereas severe post-1965 typhoons did not.[28]

[28]We note that the estimate for α_0 is - 0.742 when using ln(*male cohort size*) as the outcome variable. This estimate is much smaller than the corresponding coefficient estimate, -0.144, in Table 3 (column 2). The treatment variable in Table 3 measures the expected number of in utero typhoon exposures for each birth cohort and weighs each typhoon that took place one year before or during the birth-year by the probability that the typhoon took place in utero for a given birth cohort. The treatment variable here is an indicator of whether any typhoon passed by either one year before or during the year of birth – it does not exclude typhoons that took place pre-conception; nor does it place a lower weight on typhoons that took place close to the end of the birth-year. The estimated effects here are the combined effects of in utero, pre-conception, and post-natal exposures and, hence, smaller than the estimates in Table 3. We also used the 1970 Census and find qualitatively similar results on mortality among those exposed to severe typhoons pre- and post-1965.

Interestingly, however, α_{-1} and β_{-1}, which measure the effects of severe typhoons that took place before conception (two to three years before birth), are significantly different from zero for some outcome variables. Our results suggest that pre-conception exposure to severe, pre-1965 typhoons reduces *male* cohort size by 4 percent and the fraction of males by 0.893 percentage points, but has no detectable effects on *female* cohort size. For severe post-1965 typhoons, our results suggest that pre-conception exposure reduces *male* cohort size by 2.3 percent and reduces *female* cohort size by 4 percent. The effects on the fraction of males are small and not statistically significant. One channel through which pre-conception typhoon exposure can affect cohort size (and long-term human outcomes) is reduced household consumption and nutrient intake.[29]

Next, we perform the same analysis on long-term human capital outcomes. We estimate the equivalent of Equation 4 for education and occupational outcomes at the individual level and include as a control variable a dummy for male. Figure 2 presents the results (Estimates are presented in Appendix Table A.9). Again, pre-1965 typhoons had little impact on long-term outcomes, whereas post-1965 typhoons had large negative effects on both educational attainment and occupational skill level. In utero exposure has the largest impact; post-natal early childhood exposure has smaller but substantial impacts as well.

Alternative cutoff year Although Marcos came into office after December 1965, the coordinating body to administer disaster relief was only established in December 1968 after the Casiguran earthquake. For robustness, we use December 1968 as the alternative cut-off

However, the 1970 census provides the province of birth, not the municipality of birth. The analysis was, thus, done at the province level and the estimates are not precisely measured. These results are available upon request.

[29]We concede that it is not clear why pre-conception exposure to pre-1965 typhoons disproportionally affects male cohort size, whereas pre-conception exposure to post-1965 typhoons disproportionally affects female cohort size. We first note that, traditionally, most ethnic groups in the Philippines do not practice son preference. Prior research suggest that reductions in the nutrient intake may explain these findings. Mathews et al. (2008) provide some evidence that higher levels of maternal nutrition prior to conception is associated with a higher likelihood of male birth. Our pre-1965 results suggest that mothers who were exposed to pre-1965 typhoons prior to conception may have reduced their pre-conception nutrient intake, which lowered sex ratio. Anttila-Hughes and Hsiang (2013) find reductions in household assets and consumption as well as an increase in female infant mortality rate *three years after* typhoon exposure. Our post-1965 results are largely consistent with their findings.

time for policy change (Appendix Tables A.10, A.11, and A.12). We find similar patterns and an even greater contrast between the two time periods. Effects on cohort sizes are larger for the pre-Marcos period and smaller for the period under Marcos. Effects for most long-term outcomes remain small and insignificant pre-Marcos, whereas long-term effects for the period under Marcos become even larger. Similar to our earlier results, when we use the adult co-resident sample, the estimated effects are slightly larger with household fixed effects and for low-SES households. These results suggest that the change in the availability of disaster relief funding after December 1968 is the main contributor to the muted mortality effects in the period under Marcos.

Alternative exposure variables We use the expected number of typhoons an individual is exposed to as our main treatment variable. We assume a linear relationship between the expected number of typhoons in each stage of life and the outcome variables of interest. If some municipalities are exposed to multiple typhoons within a year and if there is a non-linear relationship between the number of exposures and the effects of each exposure, then Equations 1, 2, and 3 would lead to biased estimates. The concern for multiple typhoons is minimal – all municipalities are exposed to at most one severe typhoon each year (Table 1); in less than 1 percent of municipality-year pairs, the municipality was struck by multiple small typhoons within the same year (mostly category 1).

To further alleviate the non-linearity concerns, we expand the analysis in Equation 4 to include exposure dummies for small typhoons. We also replace the exposure dummies in Equation 4 with count variables indicating the number of typhoon exposures during each stage of life. These two sets of alternative treatment variables (exposure dummies and count variables) yield very similar results, confirming that non-linearity concerns are minimal. These results are available upon request.

Alternative distance to the eye of the storm One limitation of our study is that we do not observe the actual size of the typhoon. In our main analysis, we assume that only

municipalities within 30 kilometers of the eye of the storm are affected by the typhoon. If a typhoon is particularly large in size, municipalities outside of the 30-km radius may be just as affected as those within the 30-km radius. If this is the case for some typhoons, we may under-estimate the adverse effects of typhoons.[30]

To test whether the effects of typhoons extend beyond the 30-km radius, we add a second layer of treatment variables to municipalities that are between 30 and 60 kilometers from the eye of the storm and estimate the effects of typhoon exposure on them. These results are presented in Section 7.7 of the Appendix. Pre-1965 typhoons have limited effects on municipalities farther away from the eye of the storm. However, post-1965 typhoons have some negative effects on the cohort size of municipalities 30 to 60 kilometers from the storm path. The average size of post-1965 severe typhoons may be larger than pre-1965 severe typhoons, or disaster relief funding may have been more available for municipalities closer to the typhoon path.

Alternative storm intensity measures Our main analysis allows for only two storm intensity levels (small and severe). In Appendix Section 7.8, we allow for four levels: SS scale 1, 2, 3, and 4 or higher. We continue to combine SS scale 4 and 5 storms in one category because only a small number of municipalities were exposed to scale 5 storms. The results suggest that (1) generally, the impact of the storm increases with storm intensity, and (2) the magnitudes of the adverse effects are much larger for scale 4 and 5 storms than for scale 1, 2, or 3 storms. These results support our categorization of storm intensity.

Alternative explanations for the changing effects after 1965 Economic growth could have contributed to the changing effects of severe typhoons under the Marcos regime. However, the country's rapid growth began after 1970; growth rate was approximately con-

[30]Brand and Blelloch (1973) document that intense typhoons passing the Philippines between 1960 to 1970 have an average *eye diameter* of 20 to 30 miles, which translate to an *eye radius* of 16 to 24 kilometers. They also found that intense typhoons have smaller eye diameters but larger circulation sizes than less intense typhoons.

stant between 1960 and 1970.[31] Hence, if economic development drives the changes in the mortality effects, we would expect the changes in the effects of typhoons to begin either before 1965 or after 1970. To explore the potential impact of economic development, we separate pre-1965 severe typhoons by five-year windows to estimate the effects of severe typhoons over time (Appendix Tables A.17 and A.18). Our results suggest that neither the mortality effects nor the long-term effects changed much between 1956 -1960 and 1961-1965. In contrast, there is a sharp change in the effects of severe typhoons after 1965. Given that growth rate in the Philippines was approximately constant between 1960 and 1970, economic development is unlikely to be the cause of the sharp change after 1965.

To the best of our knowledge, the post-disaster relief policy initiated by the Marcos regime is the only typhoon-related policy change between 1965 and 1972. Some health policies that might alleviate the deleterious effects of typhoons, such as the Expanded Program on Immunization, were implemented after our period of interest[32]. Others, such as the Rural Health Unit program and the establishment of a water and sanitation regulating body, were implemented in the 1950s, long before Marcos came to power. Although we are unable to rule out other factors that may have contributed to the post-1965 changes in the effects of severe typhoons, all our results are consistent with the hypothesis that the post-disaster relief policy provided protection against typhoon-induced early-life mortality.

Migration and Immigration Migration is common in the Philippines (Warren, 2013; Abad, 1981); however, the share of the population that ever migrated internally remains fairly constant between 1960 and 1970 (Abad, 1981); international migration began increasing in the 1970s[33], after our period of interest. More importantly, migration, whether internal or international, is unlikely to affect our outcomes of interest, because (1) all individuals,

[31] Average annual per capita GDP growth rate in the Philippines was 1.9% between 1961 and 1965, 1.6% between 1966 and 1970, 3% in the 1970s, -0.9% in the 1980s, 0.5% in the 1990s and 3.4% between 2001 and 2016 (author's calculations with data from the World Bank's World Development Indicators).

[32] The immunization program targets early-life mortality directly. This was implemented in 1976. www.doh.gov.ph/expanded-program-on-immunization

[33] https://dirp4.pids.gov.ph/ris/dps/pidsdps0933.pdf; Last accessed July 24, 2018

including overseas workers, are included in the CPH 1990,[34] and (2) we observe each individual's municipality of birth, measure cohort size by birth place, and assign early-life typhoon exposure by birth place.

However, early-life migration shortly after birth may introduce measurement error in our *post*-utero exposure variables since we assign the probability of exposure in the first two years of life based on individuals' municipality of birth, but we are unable to determine individuals' municipality of residence in their first or second year of life. We explore the extent of early-life migration in the Appendix using CPH 1990 and the 1993 Philippines Demographic and Health Survey (1993 DHS). In both samples, we find that the probability of migration among households with children under 5 is low.

6 Conclusion

We find a strong dose-response effect to early-life exposure to natural disasters in a setting where natural disasters occur frequently – severe disasters are associated with adverse outcomes, whereas less intense events are associated with small or insignificant effects. We also find a strong negative relationship between mortality and long-term scarring effects. When the mortality effects of severe disasters are especially high (pre-1965 in our setting), survivors exhibited similar long-term outcomes as those who were not exposed to the shock. This is consistent with findings in high selective mortality settings. After an increase in disaster relief efforts (post-1965 in our setting), the mortality effects of severe disasters are much more muted, and we observe large differences in the long-term outcomes for the survivors and the unaffected. The observed adverse outcomes due to scarring in a low mortality setting also reflect improved early-life survival. These contrasts suggest that research on early-life shocks in developing countries should pay special attention to selective mortality, since ob-

[34]The CPH 1990 manual contains the following definition: household members include family members who are overseas workers and who are away at the time of the census are considered members of the household even though they are expected to be away for more than a year.. Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D020.V7.0

served adverse long-term outcomes may be the result of the increased probability of survival (Currie and Vogl, 2013).

The provision of resources in the aftermath of a natural disaster has long been the focus in policy making in many countries and our findings underscore the importance of such assistance. Residents, particularly low-SES ones, are not prepared to cope with severe disasters on their own. Therefore, policy makers should take into account both the community's familiarity with the disaster and the severity of natural disasters in implementing post-disaster interventions. Our results suggest that short-term assistance has been especially effective in lowering the rate of early-life mortality caused by disasters. However, alleviating the long-term effects remains a challenge for future research and policy making. To this end, our finding that children from high-SES families in the post-1965 sample were somewhat shielded from the negative effects of severe typhoons offers a sense of hope – with strong infrastructure and sufficient post-disaster aid, complete resilience may be within reach even after the most ferocious disasters.

References

Abad, R. G. (1981). Internal migration in the philippines: A review of research findings. *Philippine Studies 29*(2), 129–143.

Adhvaryu, A., T. Molina, A. Nyshadham, and J. Tamayo (2015). Helping children catch up: Early life shocks and the Progresa experiment.

Aguilar, A. and M. Vicarelli (2011). El Nino and Mexican children: Medium-term effects of early-life weather shocks on cognitive and health outcomes. *Cambridge, United States: Harvard University, Department of Economics. Manuscript*.

Almond, D. (2006). Is the 1918 Influenza pandemic over? Long-term effects of in utero Influenza exposure in the post-1940 US population. *Journal of Political Economy 114*(4), 672–712.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *The Journal of Economic Perspectives 25*(3), 153–172.

Almond, D., J. Currie, and V. Duque (2017). Childhood circumstances and adult outcomes: Act II. Technical report, National Bureau of Economic Research.

Almond, D., L. Edlund, and M. Palme (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *The Quarterly Journal of Economics 124*(4), 1729–1772.

Almond, D. and B. Mazumder (2011). Health capital and the prenatal environment: the effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics 3*(4), 56–85.

Almond, D. and B. Mazumder (2013). Fetal origins and parental responses. *Annu. Rev. Econ. 5*(1), 37–56.

Anttila-Hughes, J. K. and S. M. Hsiang (2013). Destruction, disinvestment, and death: Economic and human losses following environmental disaster. *Available at SSRN 2220501*.

Arceo, E., R. Hanna, and P. Oliva (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *The Economic Journal*.

Atkinson, J., A. Hicken, and N. Ravanilla (2014). Pork and typhoons: The influence of political connections on disaster response.

Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ: British Medical Journal 311*(6998), 171.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque (2009). Adult height and childhood disease. *Demography 46*(4), 647–669.

Brand, S. and J. W. Blelloch (1973). Changes in the characteristics of typhoons crossing the Philippines. *Journal of Applied Meteorology 12*(1), 104–109.

Cunha, F. and J. Heckman (2007). The economics of human development: The technology of skill formation. *The American Economic Review 97*(2), 31–47.

Currie, J. and D. Almond (2011). Human capital development before age five. *Handbook of Labor Economics 4*, 1315–1486.

Currie, J. and M. Rossin-Slater (2013). Weathering the storm: Hurricanes and birth outcomes. *Journal of Health Economics 32*(3), 487–503.

Currie, J. and T. Vogl (2013). Early-life health and adult circumstance in developing countries. *Annu. Rev. Econ 5*, 1–36.

de Oliveira, V. H. and C. Quintana-Domeque (2016). Natural disasters and early human development: Hurricane Catarina and birth outcomes in Brazil.

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences 104*(33), 13232–13237.

Deuchert, E. and C. Felfe (2015). The tempest: Short-and long-term consequences of a natural disaster for children's development. *European Economic Review 80*, 280–294.

Doocy, S., A. Rofi, C. Moodie, E. Spring, S. Bradley, G. Burnham, and C. Robinson (2007). Tsunami mortality in Aceh province, Indonesia. *Bulletin of the World Health Organization 85*(4), 273–278.

Franklin, S. and J. Labonne (2017). Economic shocks and labour market flexibility. *Journal of Human Resources*, 0616–8012r1.

Gunnsteinsson, S., A. Adhvaryu, P. Christian, A. Labrique, J. Sugimoto, A. A. Shamim, and K. P. West Jr (2016). Resilience to early life shocks: Evidence from the interaction of a natural experiment and a randomized control trial.

Hanna, R. and P. Oliva (2016). Implications of Climate Change for Children in Developing Countries. *The Future of Children*, 115–132.

Hsiang, S. M. and A. S. Jina (2014). The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones.

Huigen, M. G. and I. C. Jens (2006). Socio-economic impact of super typhoon Harurot in San Mariano, Isabela, the Philippines. *World Development 34*(12), 2116–2136.

Jayachandran, S. (2009). Air quality and early-life mortality evidence from Indonesia's wildfires. *Journal of Human Resources 44*(4), 916–954.

Kannisto, V., K. Christensen, and J. W. Vaupel (1997). No increased mortality in later life for cohorts born during famine. *American Journal of Epidemiology 145*(11), 987–994.

Karbownik, K. and A. Wray (2016). Long-run consequences of exposure to natural disasters.

Knapp, K. R., J. A. Knaff, C. R. Sampson, G. M. Riggio, and A. D. Schnapp (2013). A pressure-based analysis of the historical western North Pacific tropical cyclone intensity record. *Monthly Weather Review 141*(8), 2611–2631.

Kraemer, S. (2000). The fragile male. *British Medical Journal 321*(7276), 1609.

Lam, H., M. H. Kok, and K. K. Y. Shum (2012). Benefits from typhoons–the Hong Kong perspective. *Weather 67*(1), 16–21.

Lin, M.-J. and E. M. Liu (2014). Does in utero exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. *Journal of Health Economics 37*, 152–163.

Liu, E., J.-T. Liu, and T.-Y. H. Tseng (2016). Fetal origins of mental health: Evidence from natural disasters in Taiwan.

Mathews, F., P. J. Johnson, and A. Neil (2008). You are what your mother eats: evidence for maternal preconception diet influencing foetal sex in humans. *Proceedings of the Royal Society of London B: Biological Sciences 275*(1643), 1661–1668.

Mathews, T., B. E. Hamilton, et al. (2005). Trend analysis of the sex ratio at birth in the United States. *National vital statistics reports 53*(20), 1–17.

Meng, X. and N. Qian (2009). The long term consequences of famine on survivors: Evidence from a unique natural experiment using China's great famine. Technical report, National Bureau of Economic Research.

Simeonova, E. (2011). Out of sight, out of mind? Natural disasters and pregnancy outcomes in the USA. *Cesifo Economic Studies 57*(3), 403–431(29).

Ugaz, J. and A. Zanolini (2011). Effects of extreme weather shocks during pregnancy and early life on later health outcomes: the case of Philippines' typhoons.

Warren, J. F. (2013). A tale of two decades: Typhoons and floods, Manila and the Provinces, and the Marcos years. *The Asia-Pacific Journal 11*, 43.

Yu, J., D. Tang, Y. Li, Z. Huang, and G. Chen (2013). Increase in fish abundance during two typhoons in the South China Sea. *Advances in Space Research 51*(9), 1734–1749.

Zivin, J. S. G., S. M. Hsiang, and M. J. Neidell (2018). Temperature and human capital in the short and long run. *Journal of the Association of Environmental and Resource Economists 5*(1), 77–105.

Tables and Figures

Table 1: Average Typhoon Exposures Across the Philippines, 1945-1972

Panel A: Average number of typhoons affecting **the Philippines** per year

Time Period	All Typhoons	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Small	Severe
Period 1: 1945-1965	4.95	2.86	0.62	0.90	0.52	0.05	4.38	0.57
Period 2: 1966-1972	6.43	3.14	1.42	1.00	0.43	0.43	5.57	0.86
Overall: 1945-1972	5.32	2.93	0.82	0.93	0.50	0.14	4.68	0.64

Panel B: Average number of typhoons affecting **each municipality** per year, by island group

		Period 1: 1945-1965			Period 2: 1966-1972		
Island Group	Storm Intensity	Mean	Std. Dev.	Max	Mean	Std. Dev.	Max
Northern Luzon	Small	0.320	(0.561)	3	0.295	(0.556)	2
	Severe	0.010	(0.099)	1	0.019	(0.137)	1
Southern Luzon	Small	0.187	(0.430)	2	0.383	(0.641)	4
	Severe	0.001	(0.033)	1	0.049	(0.216)	1
Visayas	Small	0.155	(0.394)	3	0.259	(0.507)	3
	Severe	0.015	(0.123)	1	0.005	(0.072)	1
Mindanao	Small	0.030	(0.179)	2	0.078	(0.268)	1
	Severe	0.002	(0.044)	1	0	(0)	0
Overall	Small	0.172	(0.426)	3	0.248	(0.517)	4
	Severe	0.007	(0.085)	1	0.017	(0.130)	1

Note: Authors' calculations using JMA and TD-9635 data. Panel A shows the average number of typhoons that cross the Philippine archipelago every year. Panel B shows the average number of typhoons affecting each municipality per year. Geographic divisions and municipality boundaries are consistent with the 1990 Philippine Census. Number of municipalities = 1611. A municipality is affected if the centroid of the municipality lies within 30 km of the typhoon path. Typhoon intensity in Panel A refers to the highest intensity a typhoon ever reached over the Philippine archipelago. Storm intensity in Panel B refers to the intensity when the typhoon passed the corresponding municipality. A small typhoon is one whose highest intensity is SS scale 1, 2, or 3. A severe typhoon is one whose highest intensity is SS scale 4 or 5.

Table 2: Effects on Cohort Size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: Ages Cohorts 2 to 43				**Sample: Ages Cohorts 18 to 43**			
	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male
Small typhoon in utero	-0.00841	-0.00612	-0.0138**	0.00162	-0.0110	-0.00749	-0.0174*	0.00181
	(0.00532)	(0.00639)	(0.00676)	(0.00169)	(0.00706)	(0.00888)	(0.00939)	(0.00252)
Small typhoon year 1&2	0.00102	-0.00274	0.00431	-0.00175**	-0.000798	-0.00209	5.23e-07	-0.000491
	(0.00265)	(0.00318)	(0.00308)	(0.000721)	(0.00326)	(0.00427)	(0.00390)	(0.00107)
Severe typhoon in utero	-0.0502**	-0.0769**	-0.0267	-0.0128*	-0.0735**	-0.111**	-0.0437	-0.0175
	(0.0253)	(0.0329)	(0.0296)	(0.00760)	(0.0304)	(0.0450)	(0.0363)	(0.0109)
Severe typhoon year 1&2	0.00919	0.00121	0.0156	-0.00297	0.00751	0.00375	0.00781	-0.000192
	(0.0115)	(0.0146)	(0.0129)	(0.00337)	(0.0128)	(0.0191)	(0.0147)	(0.00504)
Observations	62,286	62,286	62,286	62,286	38,558	38,558	38,558	38,558
R-squared	0.943	0.910	0.906	0.033	0.939	0.896	0.895	0.045
Mean of Cohort Size	724.68	365.78	358.91	0.5078	576.69	287.98	288.71	0.5056

Notes: Source for the outcome variables is the 10% housing sample of the CPH 1990. Cohort size is the estimated size of each cohort, estimated by summing up the weights of all individuals with non-missing information about the municipality of birth. Each column is a separate regression. Regressions are run at the birth-municipality by age-cohort level. For columns (1) to (4), the sample includes all cohorts aged 2 to 43 in 1990. For columns (5) to (8), the sample includes all cohorts aged 18 to 43 in 1990. For all columns, municipalities are restricted to those that have at least one male and one female in each age cohort for all ages under 43.

Small typhoon in utero is the expected number of small typhoons that passed within 30 km of the cohort's municipality of birth when the cohort was in utero. Similarly, small typhoon in 1st and 2nd years are the expected number of small typhoon during the first and second years after birth. Small typhoons are those with minimum central pressure between 945 and 999 mb, which correspond to a category 1, 2, or 3 typhoon on the Saffir-Simpson scale. Severe typhoons are those with central pressure at or below 944 mb, which correspond to category 4 and 5 typhoons on the Saffir-Simpson scale.

All regressions include municipality fixed effects, birth-year by island group fixed effects, and region-specific time trends. Standard errors are clustered two-way at both the municipality and the province-by-birth-year levels.

Table 3: Effects on Cohort Size - Before and After 1965

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: Ages Cohorts 2 to 43				**Sample: Ages Cohorts 18 to 43**			
	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male
Small typhoon in utero	-0.00746	-0.00862	-0.00962	-0.000328	-0.0113	-0.0120	-0.0133	-0.000291
pre-1965	(0.00959)	(0.0119)	(0.0119)	(0.00303)	(0.00864)	(0.0113)	(0.0108)	(0.00302)
Small typhoon year 1&2	0.00883*	0.00640	0.0101*	-0.000885	0.00643	0.00440	0.00773	-0.000812
pre-1965	(0.00502)	(0.00609)	(0.00569)	(0.00129)	(0.00432)	(0.00563)	(0.00499)	(0.00133)
Severe typhoon in utero	-0.0954***	-0.144***	-0.0602	-0.0224	-0.0871**	-0.133**	-0.0517	-0.0217
pre-1965	(0.0354)	(0.0552)	(0.0456)	(0.0144)	(0.0354)	(0.0560)	(0.0470)	(0.0147)
Severe typhoon year 1&2	0.00874	0.00423	0.00722	0.000590	0.00771	0.00816	0.00385	0.00238
pre-1965	(0.0186)	(0.0282)	(0.0181)	(0.00642)	(0.0168)	(0.0269)	(0.0168)	(0.00673)
Small typhoon in utero	-0.00944	-0.00499	-0.0169**	0.00287	-0.0126	4.32e-05	-0.0281	0.00627
post-1965	(0.00624)	(0.00735)	(0.00803)	(0.00198)	(0.0122)	(0.0149)	(0.0172)	(0.00444)
Small typhoon year 1&2	-0.00361	-0.00807**	0.000881	-0.00224***	-0.0147***	-0.0142**	-0.0151**	0.000283
post-1965	(0.00312)	(0.00356)	(0.00370)	(0.000820)	(0.00499)	(0.00635)	(0.00619)	(0.00169)
Severe typhoon in utero	-0.0183	-0.0287	-0.00410	-0.00556	-0.0444	-0.0584	-0.0309	-0.00598
post-1965	(0.0359)	(0.0405)	(0.0380)	(0.00766)	(0.0528)	(0.0700)	(0.0545)	(0.0152)
Severe typhoon year 1&2	0.00740	-0.00350	0.0192	-0.00550	-0.00161	-0.0126	0.00582	-0.00470
post-1965	(0.0146)	(0.0150)	(0.0182)	(0.00369)	(0.0185)	(0.0207)	(0.0273)	(0.00693)
Observations	62,286	62,286	62,286	62,286	38,558	38,558	38,558	38,558
R-squared	0.943	0.910	0.906	0.033	0.939	0.896	0.895	0.045
Mean of Y, pre-1965	507.53	254.53	253.01	0.5057	507.53	254.53	253.01	0.5057
Mean of Y, post-1965	904.07	457.68	446.39	0.5096	764.40	378.79	385.62	0.5054

Notes: The typhoon exposure variables are interacted with a dummy variable indicating whether the typhoon occurred before or after December 1965. All regressions include municipality fixed effects, birth-year by island group fixed effects, and region-specific time trends. Standard errors are clustered two-way at both the municipality and the province-by-birth-year levels. The mean of each outcome variable is presented in the last two rows. "Mean of Y, pre-1965" refers to the averages for cohorts aged 25 to 43. "Mean of Y, post-1965" refers to the averages for cohorts aged 2 to 24 for columns (1) to (4) and cohorts aged 18 to 24 for columns (5) to (8).

Table 4: Effects on Educational Attainment

	(1)	(2)	(3)
	Literacy	Years of Education	Completed High Sch.
Small typhoon in utero	0.00129	0.0287*	0.000422
pre-1965	(0.000846)	(0.0173)	(0.00232)
Small typhoon year 1&2	-7.12e-05	0.00735	-0.000707
pre-1965	(0.000437)	(0.00911)	(0.00124)
Severe typhoon in utero	0.00466	-0.0300	-0.0143
pre-1965	(0.00407)	(0.0666)	(0.00874)
Severe typhoon year 1&2	0.00145	-0.0595	-0.0143***
pre-1965	(0.00209)	(0.0384)	(0.00529)
Small typhoon in utero	-0.00145	-0.0119	-0.000968
post-1965	(0.00116)	(0.0228)	(0.00286)
Small typhoon year 1&2	-0.000635	-0.00669	-0.00159
post-1965	(0.000603)	(0.0135)	(0.00155)
Severe typhoon in utero	-0.00771*	-0.342***	-0.0181
post-1965	(0.00394)	(0.109)	(0.0139)
Severe typhoon year 1&2	-0.000442	-0.0614	-0.00966
post-1965	(0.00184)	(0.0465)	(0.00651)
Observations	2,290,886	2,255,017	2,255,017
R-squared	0.143	0.188	0.135
Mean of Y, pre-1965 cohorts	0.950	9.19	0.444
Mean of Y, post-1965 cohorts	0.961	9.70	0.519

Notes: Sample includes all individuals between the ages of 18 and 43 with non-missing information about the municipality of birth in the 10% housing sample of CPH1990. "Literacy" is a dummy variable indicating whether the respondent was literate as of May, 1990. "Years of education" refers to the respondent's completed years of education as of May, 1990. "Completed High School" is a dummy variable indicating whether the correspondent had completed high school as of May, 1990. Each column is a separate regression. Regressions are run at the individual level, whereas the treatment variables are defined at the birth-municipality by age-cohort level. Definitions of treatment variables are the same as in Table 4. All regressions include municipality fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table 5: Effects on Occupational Skill Level

	(1) Skilled Occupation	(2) Associate Professional	(3) Professional	(4) Skilled Occupation	(5) Associate Professional	(6) Professional
Small typhoon in utero	0.00933***	0.00321**	0.00274**	0.00831***	0.00272*	0.00222*
pre-1965	(0.00299)	(0.00154)	(0.00131)	(0.00280)	(0.00139)	(0.00120)
Small typhoon year 1&2	0.00637***	0.00332***	0.00259***	0.00608***	0.00316***	0.00245***
pre-1965	(0.00172)	(0.00107)	(0.000921)	(0.00158)	(0.000998)	(0.000867)
Severe typhoon in utero	-0.00302	0.00949**	0.00794*	-0.00397	0.00871*	0.00700
pre-1965	(0.0105)	(0.00470)	(0.00462)	(0.0105)	(0.00454)	(0.00458)
Severe typhoon year 1&2	0.00108	0.00103	-0.000674	0.00290	0.00232	0.000607
pre-1965	(0.00550)	(0.00215)	(0.00215)	(0.00555)	(0.00206)	(0.00172)
Small typhoon in utero	0.00930	0.00217	0.00206	0.00976	0.00230	0.00224
post-1965	(0.00742)	(0.00274)	(0.00212)	(0.00711)	(0.00247)	(0.00189)
Small typhoon year 1&2	0.00438	0.00150	0.00154	0.00470	0.00180	0.00182
post-1965	(0.00330)	(0.00173)	(0.00147)	(0.00307)	(0.00157)	(0.00134)
Severe typhoon in utero	-0.159***	-0.0690***	-0.0519***	-0.145***	-0.0601***	-0.0448***
post-1965	(0.0421)	(0.0176)	(0.0131)	(0.0397)	(0.0157)	(0.0115)
Severe typhoon year 1&2	-0.0508***	-0.0252***	-0.0206***	-0.0500***	-0.0244***	-0.0198***
post-1965	(0.0113)	(0.00501)	(0.00392)	(0.0111)	(0.00492)	(0.00389)
Years of education	No	No	No	0.0381***	0.0252***	0.0215***
				(0.000665)	(0.000711)	(0.000615)
Observations	2,093,804	2,093,804	2,093,804	2,069,113	2,069,113	2,069,113
R-squared	0.146	0.045	0.037	0.232	0.146	0.125
Mean of Y, pre-1965 cohorts	0.313	0.101	0.084	0.313	0.101	0.084
Mean of Y, post-1965 cohorts	0.172	0.035	0.027	0.172	0.035	0.027

Notes: Outcome variables are dummy variables indicating whether the individual holds a skilled occupation (columns 1 and 4), an associate professional occupation (columns 2 and 5), or a professional occupation (columns 3 and 6). Each column is a separate regression. Regressions are run at the individual level, whereas the treatment variables are defined at the birth-municipality by age-cohort level. All regressions include municipality fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table 6: Effects on Educational Attainment by Sex

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy		Years of Education		Completed High Sch.	
	Male	Female	Male	Female	Male	Female
Small typhoon in utero	0.00102	0.00154	0.00733	0.0516**	-0.000860	0.00193
pre-1965	(0.00107)	(0.00114)	(0.0229)	(0.0219)	(0.00314)	(0.00293)
Small typhoon year 1&2	2.48e-05	-0.000173	0.00691	0.00779	-0.000786	-0.000575
pre-1965	(0.000546)	(0.000557)	(0.0107)	(0.0116)	(0.00145)	(0.00161)
Severe typhoon in utero	0.00440	0.00480	-0.00734	-0.0540	-0.0215*	-0.00675
pre-1965	(0.00530)	(0.00555)	(0.102)	(0.0856)	(0.0125)	(0.0109)
Severe typhoon year 1&2	0.00191	0.000707	-0.0409	-0.0790	-0.0178**	-0.0107*
pre-1965	(0.00256)	(0.00318)	(0.0562)	(0.0485)	(0.00764)	(0.00568)
Small typhoon in utero	-0.000860	-0.00209	-0.0272	0.00467	-0.00263	0.00146
post-1965	(0.00141)	(0.00145)	(0.0262)	(0.0301)	(0.00395)	(0.00356)
Small typhoon year 1&2	-0.000608	-0.000738	0.00163	-0.0150	-0.000157	-0.00293
post-1965	(0.000690)	(0.000771)	(0.0132)	(0.0169)	(0.00182)	(0.00200)
Severe typhoon in utero	-0.00878*	-0.00685	-0.400***	-0.286**	-0.0259	-0.0128
post-1965	(0.00460)	(0.00520)	(0.113)	(0.131)	(0.0175)	(0.0164)
Severe typhoon year 1&2	0.000358	-0.00133	-0.0706	-0.0574	-0.00857	-0.0117
post-1965	(0.00180)	(0.00233)	(0.0435)	(0.0618)	(0.00697)	(0.00810)
Observations	1,144,609	1,146,277	1,127,900	1,127,117	1,127,900	1,127,117
R-squared	0.117	0.177	0.186	0.194	0.137	0.139
Mean of Y, pre-1965 cohorts	0.952	0.949	9.12	9.26	0.447	0.440
Mean of Y, post-1965 cohorts	0.959	0.963	9.41	9.83	0.485	0.552

Notes: Each column is a separate regression using the sub-sample of either male or female respondents in the CPH 1990 10% sample. All regressions include municipality fixed effects, birth-year by island group fixed effects, and region-specific time trends. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table 7: Effects on Occupational Skill Level by Sex

	(1)	(2)	(3)	(4)	(5)	(6)
	Skilled Occupation		Associate Professional		Professional	
	Male	Female	Male	Female	Male	Female
Small typhoon in utero,	0.0108***	0.00869***	0.00327	0.00326*	0.00247	0.00309*
pre-1965	(0.00414)	(0.00319)	(0.00221)	(0.00177)	(0.00185)	(0.00166)
Small typhoon year 1&2,	0.00494**	0.00791***	0.00308**	0.00361***	0.00269**	0.00253**
pre-1965	(0.00229)	(0.00173)	(0.00136)	(0.00115)	(0.00112)	(0.00106)
Severe typhoon in utero,	0.00612	-0.00999	-0.00362	0.0223***	-0.00908	0.0242***
pre-1965	(0.0146)	(0.0139)	(0.00826)	(0.00606)	(0.00684)	(0.00597)
Severe typhoon year 1&2,	-0.00227	0.00434	0.00315	-0.000887	0.00256	-0.00370
pre-1965	(0.00834)	(0.00431)	(0.00419)	(0.00318)	(0.00322)	(0.00326)
Small typhoon in utero,	0.0143	0.00552	0.00118	0.00301	0.00135	0.00254
post-1965	(0.0102)	(0.00609)	(0.00329)	(0.00269)	(0.00246)	(0.00219)
Small typhoon year 1&2,	0.00587	0.00287	0.00187	0.00100	0.00200	0.000974
post-1965	(0.00437)	(0.00325)	(0.00202)	(0.00164)	(0.00171)	(0.00139)
Severe typhoon in utero,	-0.183***	-0.144***	-0.0684***	-0.0703***	-0.0460***	-0.0579***
post-1965	(0.0553)	(0.0372)	(0.0194)	(0.0170)	(0.0138)	(0.0133)
Severe typhoon year 1&2,	-0.0702***	-0.0360***	-0.0307***	-0.0205***	-0.0242***	-0.0174***
post-1965	(0.0153)	(0.00991)	(0.00663)	(0.00451)	(0.00500)	(0.00389)
Observations	1,043,359	1,050,445	1,043,359	1,050,445	1,043,359	1,050,445
R-squared	0.201	0.066	0.062	0.035	0.048	0.030
Mean of Y, pre-1965 cohorts	0.415	0.221	0.101	0.102	0.078	0.090
Mean of Y, post-1965 cohorts	0.194	0.152	0.029	0.041	0.020	0.034

Notes: Each column is a separate regression using the sub-sample of either male or female respondents in the CPH 1990 10% sample. All regressions include municipality fixed effects, birth-year by island group fixed effects, and region-specific time trends. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table 8: Effects on Educational Attainment - Sibling Comparison

	(1)	(2)	(3)	(4)	(5)	(6)
	With Household Fixed Effects			**With Municipality Fixed Effects**		
	Literacy	Years of Education	Completed High Sch.	Literacy	Years of Education	Completed High Sch.
Small typhoon in utero, pre-1965	0.00109	0.0428	-0.00617	0.00181	0.0718**	-0.00318
	(0.00182)	(0.0353)	(0.00496)	(0.00144)	(0.0332)	(0.00455)
Small typhoon year 1&2, pre-1965	-0.000502	0.00539	0.000447	-0.000870	-0.0264	-0.00785***
	(0.000968)	(0.0196)	(0.00289)	(0.000806)	(0.0176)	(0.00255)
Severe typhoon in utero, pre-1965	-0.00371	-0.0775	-0.00331	-0.00137	-0.361***	-0.0444**
	(0.00862)	(0.132)	(0.0221)	(0.00721)	(0.126)	(0.0187)
Severe typhoon year 1&2, pre-1965	-0.00109	-0.0950	-0.00617	0.00113	-0.0308	-0.00616
	(0.00410)	(0.0807)	(0.0109)	(0.00364)	(0.0727)	(0.00825)
Small typhoon in utero, post-1965	0.000591	-0.000541	0.00160	-0.000840	-0.0192	0.000472
	(0.00138)	(0.0290)	(0.00412)	(0.00114)	(0.0252)	(0.00374)
Small typhoon year 1&2, post-1965	-0.000156	0.00386	-0.00236	-0.000520	0.0102	-0.000899
	(0.000640)	(0.0140)	(0.00200)	(0.000572)	(0.0124)	(0.00191)
Severe typhoon in utero, post-1965	-0.00749	-0.472***	-0.0149	-0.00419	-0.222**	0.0166
	(0.00483)	(0.122)	(0.0154)	(0.00397)	(0.0953)	(0.0168)
Severe typhoon year 1&2, post-1965	-6.18e-05	-0.0842**	-0.00722	3.40e-05	-0.00529	0.000257
	(0.00211)	(0.0382)	(0.00659)	(0.00163)	(0.0414)	(0.00613)
Observations	586,234	575,982	575,982	586,233	577,819	577,819
R-squared	0.665	0.776	0.702	0.101	0.172	0.127
Mean of Y, pre-1965	0.968	10.69	0.629	0.968	10.69	0.629
Mean of Y, post-1965	0.973	10.36	0.605	0.973	10.36	0.605

Notes: Sample restricted to those in the 10% housing sample of CPH 1990 who still live in the same household as their parents and whose reported relationship to the household head is that of either "son" or "daughter." Regressions in columns (1) to (3) include household fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male. Regressions in columns (4) to (6) include municipality fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male.

Table 9: Effects on Occupational Skill Level - Sibling Comparison

	(1)	(2)	(3)	(4)	(5)	(6)
	With Household Fixed Effects			**With Municipality Fixed Effects**		
	Skilled Occupation	Associate Professional	Professional	Skilled Occupation	Associate Professional	Professional
Small typhoon in utero, pre-1965	0.0190**	0.0137***	0.00804*	0.0279***	0.0146***	0.0108***
	(0.00769)	(0.00473)	(0.00467)	(0.00655)	(0.00380)	(0.00344)
Small typhoon year 1&2, pre-1965	0.00607	0.00110	0.000965	0.0138***	0.00601**	0.00461*
	(0.00410)	(0.00300)	(0.00297)	(0.00372)	(0.00283)	(0.00252)
Severe typhoon in utero, pre-1965	-0.00556	-0.00665	-0.00986	0.00222	0.00478	0.00506
	(0.0324)	(0.0165)	(0.0136)	(0.0311)	(0.0106)	(0.00942)
Severe typhoon year 1&2, pre-1965	-0.0343**	-0.00542	-0.0102	0.000247	0.00950	0.00181
	(0.0157)	(0.00909)	(0.00707)	(0.0155)	(0.00948)	(0.00769)
Small typhoon in utero, post-1965	0.0112*	0.00106	0.000498	0.0108*	0.00300	0.00243
	(0.00625)	(0.00275)	(0.00220)	(0.00635)	(0.00284)	(0.00236)
Small typhoon year 1&2, post-1965	0.00469	0.00238	0.00243*	0.00565*	0.00358*	0.00333*
	(0.00320)	(0.00190)	(0.00147)	(0.00289)	(0.00206)	(0.00180)
Severe typhoon in utero, post-1965	-0.126***	-0.0646***	-0.0483***	-0.124***	-0.0620***	-0.0493***
	(0.0379)	(0.0179)	(0.0129)	(0.0328)	(0.0140)	(0.0110)
Severe typhoon year 1&2, post-1965	-0.0460***	-0.0272***	-0.0212***	-0.0461***	-0.0281***	-0.0221***
	(0.0123)	(0.00583)	(0.00497)	(0.0103)	(0.00465)	(0.00394)
Observations	483,553	483,553	483,553	502,380	502,380	502,380
R-squared	0.595	0.525	0.510	0.161	0.084	0.071
Mean of Y, pre-1965	0.407	0.143	0.117	0.407	0.143	0.117
Mean of Y, post-1965	0.187	0.042	0.032	0.187	0.042	0.032

Notes: Sample restricted to those in the 10% housing sample of CPH 1990 who still live in the same household as theirparents and whose reported relationship to the household head is that of either "son" or "daughter." Regressions in columns (1) to (3) include household fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male. Regressions in columns (4) to (6) include municipality fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male.

Table 10: Effects on Education - Sibling Comparison By Household Head Occupation

	(1)	(2)	(3)	(4)	(5)	(6)
	Sample: Household Head Unskilled Occ			**Sample: Household Head Skilled Occ**		
	Literacy	Years of Education	Completed High Sch.	Literacy	Years of Education	Completed High Sch.
Small typhoon in utero	0.00260	0.0351	-0.00714	-0.00223	0.0390	-0.00177
pre-1965	(0.00224)	(0.0434)	(0.00606)	(0.00245)	(0.0590)	(0.00771)
Small typhoon year 1&2	-0.000888	-0.00245	0.000769	0.000452	0.0146	0.000977
pre-1965	(0.00119)	(0.0224)	(0.00335)	(0.00150)	(0.0331)	(0.00480)
Severe typhoon in utero	-0.00692	-0.162	-0.0103	0.00769	0.145	0.0251
pre-1965	(0.0109)	(0.170)	(0.0280)	(0.00975)	(0.195)	(0.0323)
Severe typhoon year 1&2	-0.00219	-0.109	-0.00902	0.000766	-0.0982	0.00442
pre-1965	(0.00558)	(0.0974)	(0.0122)	(0.00616)	(0.120)	(0.0186)
Small typhoon in utero	0.000858	0.000787	0.00314	-0.000394	-0.0553	-3.64e-05
post-1965	(0.00176)	(0.0313)	(0.00480)	(0.00182)	(0.0452)	(0.00695)
Small typhoon year 1&2	8.97e-05	0.00831	-0.00335	-0.000680	-0.0176	0.000206
post-1965	(0.000855)	(0.0163)	(0.00247)	(0.000811)	(0.0212)	(0.00300)
Severe typhoon in utero	-0.00884	-0.399***	-0.0159	-0.00569	-0.284*	-0.0179
post-1965	(0.00657)	(0.139)	(0.0199)	(0.00507)	(0.156)	(0.0255)
Severe typhoon year 1&2	0.000210	-0.0707	-0.00324	-0.000701	-0.0701	-0.00618
post-1965	(0.00285)	(0.0504)	(0.00804)	(0.00258)	(0.0670)	(0.00889)
Observations	412,824	405,935	405,935	173,410	170,047	170,047
R-squared	0.679	0.769	0.695	0.510	0.735	0.662
Mean of Y, pre-1965	0.956	10.20	0.577	0.988	12.17	0.787
Mean of Y, post-1965	0.964	9.72	0.526	0.992	11.75	0.776

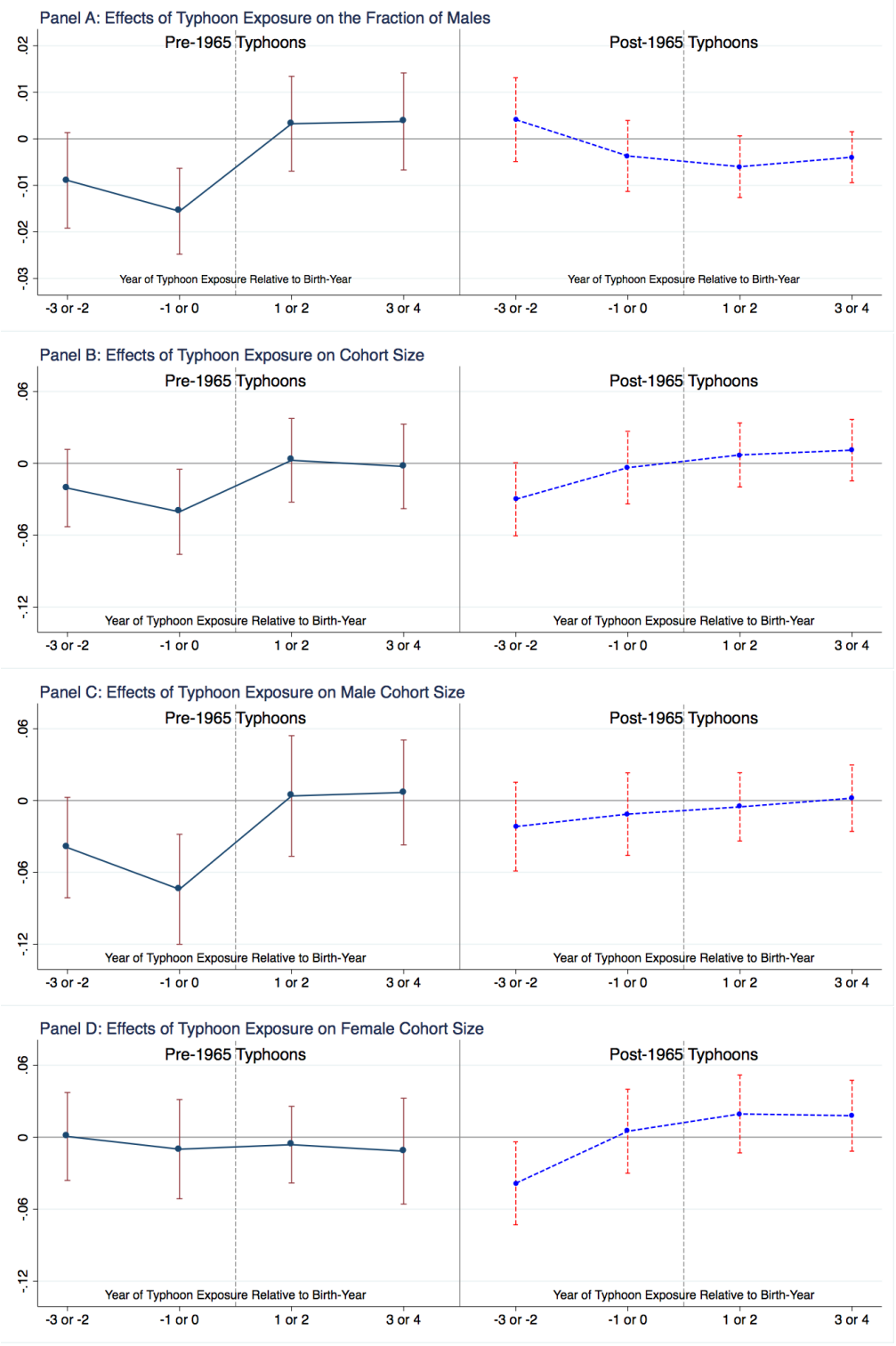
Notes: Sample restricted to those in the 10% housing sample of CPH 1990 who still live in the same household as their parents and whose reported relationship to the household head is that of either "son" or "daughter." For columns (1) to (3), the sample is further restricted to children of households where the household head has an unskilled occupation. For columns (4) to (6), the sample is further restricted to children of households where the household head has a skilled occupation. All regressions include household fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male.

Table 11: Effects on Occupation - Sibling Comparison By Household Head Occupation

	(1)	(2)	(3)	(4)	(5)	(6)
	Sample: Household Head Unskilled Occ			**Sample: Household Head Skilled Occ**		
	Skilled Occupation	Associate Professional	Professional	Skilled Occupation	Associate Professional	Professional
Small typhoon in utero, pre-1965	0.0110	0.00478	0.00151	0.0190	0.0275**	0.0184*
	(0.00856)	(0.00526)	(0.00484)	(0.0148)	(0.0119)	(0.0109)
Small typhoon year 1&2, pre-1965	0.000901	-0.00341	-0.00262	0.00717	0.00735	0.00609
	(0.00399)	(0.00284)	(0.00296)	(0.00855)	(0.00644)	(0.00613)
Severe typhoon in utero, pre-1965	-0.0341	-0.0298	-0.0216	0.0623	0.0566	0.0195
	(0.0277)	(0.0203)	(0.0177)	(0.0454)	(0.0394)	(0.0295)
Severe typhoon year 1&2, pre-1965	-0.0300**	0.00232	0.00233	-0.0774***	-0.0466**	-0.0630***
	(0.0147)	(0.00884)	(0.00824)	(0.0240)	(0.0194)	(0.0180)
Small typhoon in utero, post-1965	0.00705	-0.000598	-0.000151	0.00993	0.000870	-0.000827
	(0.00554)	(0.00256)	(0.00206)	(0.00909)	(0.00522)	(0.00436)
Small typhoon year 1&2, post-1965	0.00333	0.00220	0.00199	0.00235	0.000274	0.00154
	(0.00300)	(0.00156)	(0.00127)	(0.00486)	(0.00318)	(0.00269)
Severe typhoon in utero, post-1965	-0.107***	-0.0531***	-0.0391***	-0.0569	-0.0402*	-0.0316*
	(0.0362)	(0.0161)	(0.0119)	(0.0386)	(0.0226)	(0.0167)
Severe typhoon year 1&2, post-1965	-0.0340***	-0.0176***	-0.0149***	-0.0218	-0.0204**	-0.0118
	(0.0120)	(0.00531)	(0.00457)	(0.0168)	(0.00864)	(0.00731)
Observations	345,476	345,476	345,476	138,077	138,077	138,077
R-squared	0.587	0.512	0.496	0.599	0.539	0.526
Mean of Y, pre-1965	0.345	0.114	0.094	0.602	0.234	0.192
Mean of Y, post-1965	0.151	0.029	0.023	0.266	0.071	0.054

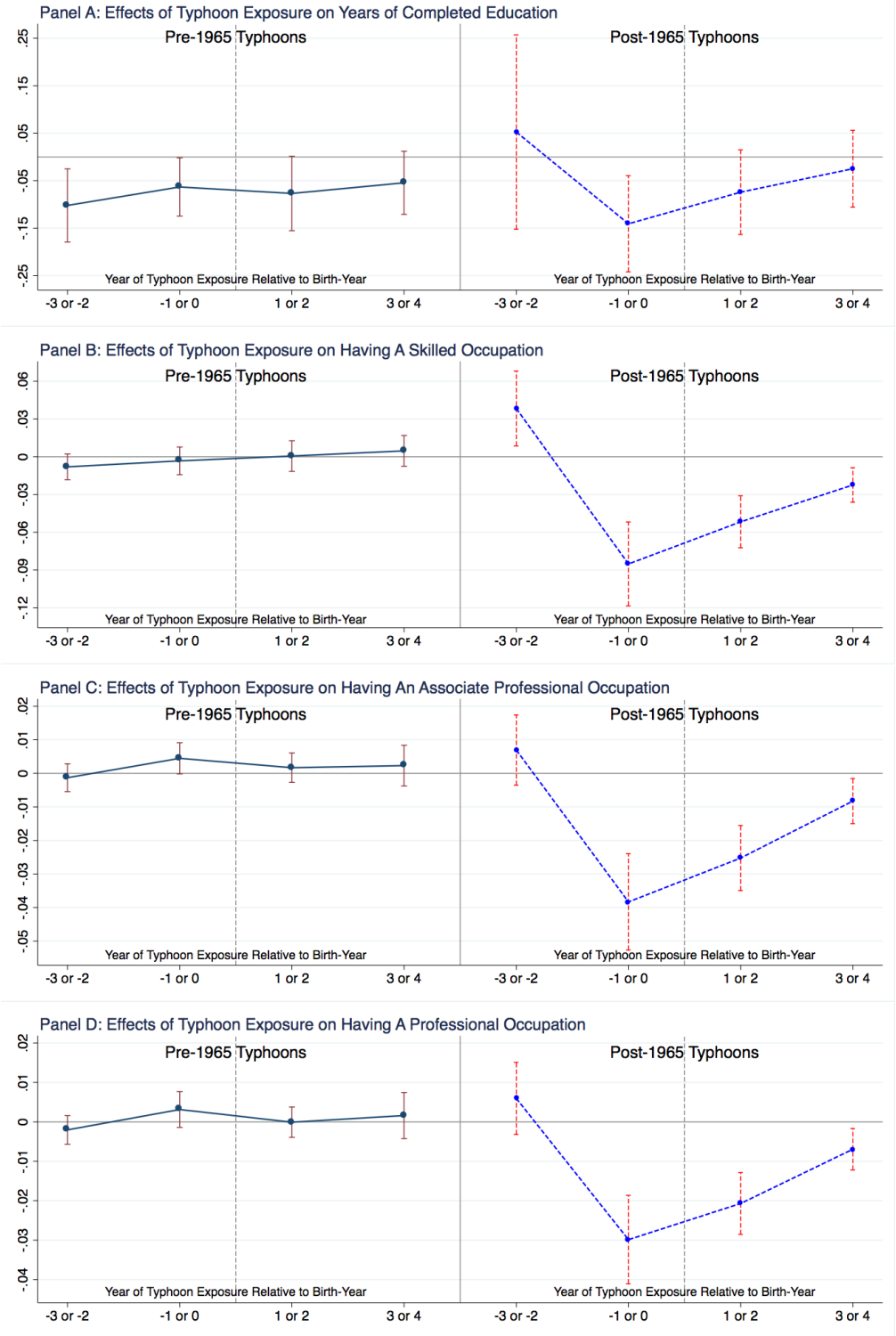
Notes: Sample restricted to those in the 10% housing sample of CPH 1990 who still live in the same household as their parents and whose reported relationship to the household head is that of either "son" or "daughter." For columns (1) to (3), the sample is further restricted to children of households where the household head has an unskilled occupation. For columns (4) to (6), the sample is further restricted to children of households where the household head has a skilled occupation. All regressions include household fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male.

Figure 1: Effects of Severe Typhoons on Cohort Size by Year of Exposure



Notes: Each panel shows the results of a separate regression where the outcome variable is regressed on eight dummy variables indicating whether a severe pre-1965 or post-1965 typhoon passed each municipality-birth-year 2 to 3 years before the birth-year / 0 to 1 year before the birth-year / 1 to 2 years after the birth-year / 3 to 4 years after the birth-year, as well as municipality fixed effects, age-by-island-group fixed effects, and region-specific time trends. The left side of each panel shows the coefficient estimates and 95% confidence intervals of the four dummy variables associated with pre-1965 typhoons, and the right side post-1965 typhoons. The regressions are run at the municipality-age-cohort level. The sample includes cohorts between the ages of 2 and 43 in 1990 and municipalities that have at least one male and one female in each age cohort for all ages under 43 in the 1990 Census 10% Housing Survey. Robust standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Figure 2: Effects of Severe Typhoons on Education and Occupation by Year of Exposure



Notes: Each panel shows the results from a separate regression where the outcome variable is regressed on eight dummy variables indicating whether a severe pre-1965 or post-1965 typhoon passed each municipality-birth-year 2 to 3 years before the birth-year / 0 to 1 year before the birth-year / 1 to 2 years after the birth-year / 3 to 4 years after the birth-year, as well as municipality fixed effects, age-by-island-group fixed effects, and region-specific time trends. The left side of each panel shows the coefficient estimates and 95% confidence intervals of the four dummy variables associated with pre-1965 typhoons, and the right side post-1965 typhoons. The sample includes all individuals between the ages of 18 and 43 with non-missing information about the municipality of birth in the 10% housing sample of CPH1990. Robust standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Online Appendix for "Selective mortality and the long-term effects of early-life exposure to natural disasters"

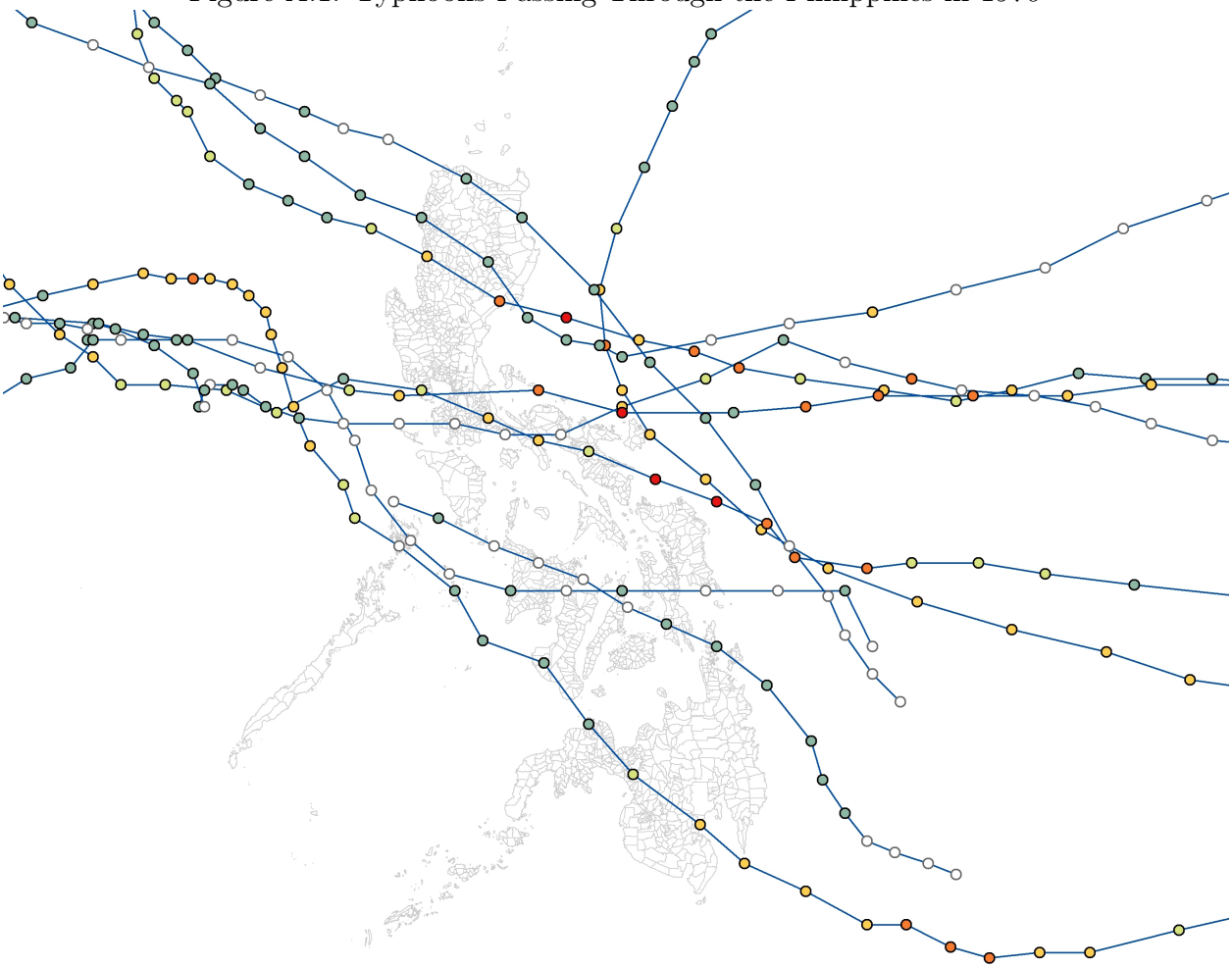
Margaret Triyana and Xing Xia

April 5, 2018

7 Supplementary Figures and Tables

7.1 Typhoon paths

Figure A.1: Typhoons Passing Through the Philippines in 1970



Notes: Authors' calculations of all typhoons that passed by the Philippines region in 1970. Each dot indicates the location and severity of the six-hourly observations of the typhoon. Red dots indicate that the typhoon was category 5 at the time it passed the location, rich orange dots category 4, yellow dots category 3, lime dots category 2, teal dots category 1, and hollow dots tropical cyclones below typhoon severity. In the background is a map of the Philippines with municipality outlines.
Data Source: Japan Meteorological Agency Tropical Cyclone Database.

7.2 Summary Statistics

Table A.1: Summary statistics

	(1) Exposed Small	(2) Exposed Severe	(3) Never Exposed
Panel A. Age 25-43, Pre-Marcos			
Literacy	0.965	0.964	0.945
	(0.183)	(0.187)	(0.228)
Years of education	9.326	9.448	9.140
	(3.747)	(3.763)	(3.943)
Completed high school	0.450	0.464	0.441
	(0.497)	(0.499)	(0.497)
Skilled occupation	0.315	0.360	0.311
	(0.464)	(0.480)	(0.463)
Associate professional occupation	0.100	0.116	0.101
	(0.300)	(0.320)	(0.302)
Professional occupation	0.083	0.095	0.084
	(0.276)	(0.293)	(0.277)
N	351,245	33,101	1,110,085
Panel B. Age 18-24, under Marcos			
Literacy	0.971	0.982	0.955
	(0.168)	(0.133)	(0.207)
Years of education	9.783	10.398	9.593
	(3.348)	(2.834)	(3.576)
Completed high school	0.526	0.620	0.505
	(0.499)	(0.486)	(0.500)
Skilled occupation	0.169	0.139	0.177
	(0.374)	(0.346)	(0.382)
Associate professional occupation	0.032	0.020	0.038
	(0.176)	(0.138)	(0.192)
Professional occupation	0.024	0.013	0.030
	(0.152)	(0.114)	(0.170)
N	214,533	61,013	551,681

Notes: Source for the outcome variables is the 10% housing sample of the CPH 1990. Outcomes for respondents who were exposed in utero to small typhoons only (col. 1), ever exposed to severe typhoons (col. 2), or never exposed to typhoons (col. 3). Small typhoon in utero is the expected number of small typhoons that passed within 30 km of the respondent's municipality of birth when the respondent was in utero. Similarly, small typhoon in the 1st and 2nd years are the expected number of small typhoon during the first and second years after birth. Small typhoons are those with minimum central pressure between 945 and 999 mb, which correspond to a category 1, 2, or 3 typhoon on the Saffir-Simpson scale. Severe typhoons are those with central pressure at or below 944 mb, which correspond to category 4 and 5 typhoons on the Saffir-Simpson scale.

7.3 Effects on Cohort Size Including All Municipalities

Table A.3: Effects on Cohort Size - All Municipalities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: Ages Cohorts 2 to 43				**Sample: Ages Cohorts 18 to 43**			
	ln(Cohort Size+1)	ln(Male Cohort+1)	ln(Female Cohort+1)	Fraction Male	ln(Cohort Size+1)	ln(Male Cohort+1)	ln(Female Cohort+1)	Fraction Male
Small typhoon in utero	-0.00728	-0.000462	-0.00984	0.00216	-0.00791	7.18e-05	-0.0113	0.00259
pre-1965	(0.0107)	(0.0134)	(0.0123)	(0.00307)	(0.00924)	(0.0123)	(0.0107)	(0.00304)
Small typhoon year 1&2	0.00650	0.00539	0.00524	-0.000136	0.00556	0.00527	0.00366	0.000156
pre-1965	(0.00570)	(0.00697)	(0.00650)	(0.00137)	(0.00490)	(0.00635)	(0.00569)	(0.00142)
Severe typhoon in utero	-0.0926**	-0.149**	-0.0492	-0.0247*	-0.0791*	-0.135**	-0.0344	-0.0248
pre-1965	(0.0464)	(0.0685)	(0.0480)	(0.0147)	(0.0419)	(0.0595)	(0.0511)	(0.0151)
Severe typhoon year 1&2	-0.00112	0.00563	-0.00441	0.00266	0.00346	0.0145	-0.00254	0.00437
pre-1965	(0.0248)	(0.0321)	(0.0251)	(0.00639)	(0.0181)	(0.0279)	(0.0180)	(0.00662)
Small typhoon in utero	-0.00817	-0.00714	-0.0142*	0.00203	-0.00771	0.000564	-0.0223	0.00594
post-1965	(0.00642)	(0.00732)	(0.00830)	(0.00200)	(0.0130)	(0.0149)	(0.0176)	(0.00437)
Small typhoon year 1&2	-0.00597*	-0.00994***	-0.00251	-0.00181**	-0.0137**	-0.0106	-0.0176***	0.00164
post-1965	(0.00322)	(0.00383)	(0.00371)	(0.000828)	(0.00572)	(0.00701)	(0.00652)	(0.00171)
Severe typhoon in utero	-0.00154	0.00169	-0.00473	0.00220	-0.0158	0.000698	-0.0392	0.00965
post-1965	(0.0377)	(0.0449)	(0.0380)	(0.00859)	(0.0581)	(0.0752)	(0.0619)	(0.0165)
Severe typhoon year 1&2	0.00589	-0.00519	0.0189	-0.00586	0.00389	-0.00981	0.0225	-0.00868
post-1965	(0.0153)	(0.0159)	(0.0190)	(0.00387)	(0.0230)	(0.0267)	(0.0299)	(0.00703)
Observations	67,578	67,578	67,578	67,554	41,834	41,834	41,834	41,810
R-squared	0.939	0.897	0.896	0.031	0.935	0.880	0.883	0.043
Mean of Y, pre-1965	475.44	238.50	236.95	0.5057	475.44	238.50	236.95	0.5057
Mean of Y, post-1965	851.09	430.97	420.12	0.5097	717.30	355.56	361.74	0.5054

Notes: Sample includes all birth-municipalities in CPH 1990. If a municipality has zero births in a given year in the 10% sample of CPH 1990, it results in a missing value for "fraction male" for the corresponding birth-municipality-birth-year. As a result, column (4) shows the results from an unbalanced panel. Other details are the same as in Table 3.

7.4 Results *Not* Separating The Eras Pre- and Post-1965

Table A.4: Effects on Educational Attainment

	(1)	(2)	(3)
	Literacy	Years of Education	Completed High Sch.
Small typhoon in utero	0.000294	0.0149	-4.65e-05
	(0.000653)	(0.0134)	(0.00177)
Small typhoon year 1&2	-0.000196	0.00370	-0.00108
	(0.000326)	(0.00735)	(0.000895)
Severe typhoon in utero	-0.00102	-0.178***	-0.0160*
	(0.00297)	(0.0658)	(0.00824)
Severe typhoon year 1&2	0.000447	-0.0593*	-0.0117***
	(0.00139)	(0.0308)	(0.00426)
Observations	2,290,886	2,255,017	2,255,017
R-squared	0.143	0.188	0.135
Mean of Y	0.954	9.36	0.468

Notes: Details are the same as in Table 4 except that typhoon variables are not interacted with pre- or post-1965 dummies.

Table A.5: Effects on Occupational Skill Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Skilled Occ.	Associate Prof.	Professional	Skilled Occ.	Associate Prof.	Professional
Small typhoon in utero	0.00989***	0.00314**	0.00271**	0.00934***	0.00282**	0.00240**
	(0.00368)	(0.00154)	(0.00121)	(0.00350)	(0.00139)	(0.00108)
Small typhoon year 1&2	0.00689***	0.00326***	0.00268***	0.00675***	0.00323***	0.00266***
	(0.00190)	(0.00120)	(0.00103)	(0.00177)	(0.00112)	(0.000976)
Severe typhoon in utero	-0.0793***	-0.0287***	-0.0212***	-0.0735***	-0.0249***	-0.0184***
	(0.0233)	(0.0101)	(0.00772)	(0.0220)	(0.00900)	(0.00684)
Severe typhoon year 1&2	-0.0259***	-0.0125***	-0.0110***	-0.0246***	-0.0115***	-0.0100***
	(0.00770)	(0.00335)	(0.00251)	(0.00765)	(0.00331)	(0.00243)
Years of education	No	No	No	0.0381***	0.0252***	0.0215***
				(0.000665)	(0.000712)	(0.000615)
Observations	2,093,804	2,093,804	2,093,804	2,069,113	2,069,113	2,069,113
R-squared	0.146	0.045	0.037	0.231	0.146	0.125
Mean of Y	0.265	0.078	0.064	0.265	0.078	0.064

Notes: Details are the same as in Table 5 except that typhoon variables are not interacted with pre- or post-1965 dummies.

Table A.6: Effects on Educational Attainment by Sex

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy		Years of Education		Completed High Sch.	
	Male	Female	Male	Female	Male	Female
Small typhoon in utero	0.000369	0.000199	-0.00436	0.0356**	-0.00155	0.00189
	(0.000839)	(0.000818)	(0.0175)	(0.0168)	(0.00251)	(0.00220)
Small typhoon year 1&2	-0.000131	-0.000296	0.00750	-9.95e-05	-0.000582	-0.00149
	(0.000404)	(0.000425)	(0.00821)	(0.00923)	(0.00108)	(0.00116)
Severe typhoon in utero	-0.00185	-0.000350	-0.199**	-0.160**	-0.0237**	-0.00938
	(0.00382)	(0.00381)	(0.0809)	(0.0783)	(0.0107)	(0.00988)
Severe typhoon year 1&2	0.00113	-0.000411	-0.0563	-0.0655	-0.0130**	-0.0109**
	(0.00154)	(0.00193)	(0.0348)	(0.0404)	(0.00513)	(0.00505)
Observations	1,144,609	1,146,277	1,127,900	1,127,117	1,127,900	1,127,117
R-squared	0.117	0.177	0.186	0.194	0.137	0.139
Mean of Y	0.955	0.954	9.20	9.51	0.458	0.479

Notes: Details are the same as in Table 6 except for that typhoon variables are not interacted with pre- or post-1965 dummies.

Table A.7: Effects on Occupational Skill Level by Sex

	(1)	(2)	(3)	(4)	(5)	(6)
	Skilled Occupation		Associate Professional		Professional	
	Male	Female	Male	Female	Male	Female
Small typhoon in utero	0.0126**	0.00817**	0.00267	0.00363**	0.00213	0.00326**
	(0.00492)	(0.00333)	(0.00191)	(0.00165)	(0.00149)	(0.00143)
Small typhoon year 1&2	0.00687***	0.00705***	0.00325**	0.00328***	0.00285**	0.00250**
	(0.00233)	(0.00194)	(0.00141)	(0.00115)	(0.00119)	(0.00102)
Severe typhoon in utero	-0.0878***	-0.0745***	-0.0351***	-0.0227**	-0.0270***	-0.0157*
	(0.0297)	(0.0216)	(0.0112)	(0.0103)	(0.00787)	(0.00877)
Severe typhoon year 1&2	-0.0377***	-0.0164**	-0.0144***	-0.0109***	-0.0114***	-0.0107***
	(0.0105)	(0.00642)	(0.00463)	(0.00307)	(0.00344)	(0.00255)
Years of education	No	No	No	No	No	No
Observations	1,043,359	1,050,445	1,043,359	1,050,445	1,043,359	1,050,445
R-squared	0.201	0.066	0.061	0.035	0.048	0.030
Mean of Y	0.339	0.191	0.076	0.080	0.058	0.070

Notes: Details are the same as in Table 7 except for that typhoon variables are not interacted with pre- or post-1965 dummies.

7.5 Effects by Year of Exposure (Regression Results for Event-Study Graphs)

Table A.8: Event Study - Effects of Severe Typhoons on Cohort Size by Year of Exposure

	(1)	(2)	(3)	(4)
	ln(Cohort Size)	ln(Male)	ln(Female)	Fraction Male
Severe typhoon pre-1965				
2 or 3 years *before* birth-year	-0.0206	-0.0393*	0.000570	-0.00893*
	(0.0165)	(0.0214)	(0.0187)	(0.00523)
0 or 1 year *before* birth-year	-0.0405**	-0.0742***	-0.0100	-0.0156***
	(0.0181)	(0.0234)	(0.0211)	(0.00471)
1 or 2 years *after* birth-year	0.00256	0.00380	-0.00626	0.00324
	(0.0178)	(0.0257)	(0.0163)	(0.00519)
3 or 4 years *after* birth-year	-0.00258	0.00675	-0.0117	0.00374
	(0.0180)	(0.0223)	(0.0225)	(0.00531)
Severe typhoon post-1965				
2 or 3 years *before* birth-year	-0.0309**	-0.0228	-0.0392**	0.00406
	(0.0155)	(0.0189)	(0.0176)	(0.00460)
0 or 1 year *before* birth-year	-0.00340	-0.0109	0.00483	-0.00354
	(0.0155)	(0.0176)	(0.0178)	(0.00388)
1 or 2 years *after* birth-year	0.00740	-0.00478	0.0195	-0.00590*
	(0.0136)	(0.0146)	(0.0165)	(0.00338)
3 or 4 years *after* birth-year	0.0112	0.00176	0.0184	-0.00413
	(0.0131)	(0.0142)	(0.0151)	(0.00280)
Observations	62,286	62,286	62,286	62,286
R-squared	0.943	0.910	0.906	0.034

Notes: Each column shows the results from a separate regression where the outcome variable is regressed on eight dummy variables indicating whether a severe pre-1965 or post-1965 typhoon passed each municipality-birth-year 2 to 3 years before the birth-year / 0 to 1 year before the birth-year / 1 to 2 years after the birth-year / 3 to 4 years after the birth-year, as well as municipality fixed effects, age-by-island-group fixed effects, and region-specific time trends. The regressions are run at the municipality-age-cohort level. Data source is the 1990 Census of the Philippines 10% Housing Survey. Sample includes cohorts between the ages of 2 and 43 in 1990 and municipalities with at least one male and one female in each age cohort for all ages under 43 in the 1990 Census 10% Housing Survey. Robust standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table A.9: Effects of Severe Typhoons on Education and Occupation by Year of Exposure

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy	Years of Education	Completed High Sch.	Skilled Occupation	Associate Professional	Professional
Severe typhoon pre-1965						
2 or 3 years *before* birth-year	0.00257	-0.102***	-0.0167***	-0.00801	-0.00129	-0.00203
	(0.00183)	(0.0393)	(0.00543)	(0.00522)	(0.00211)	(0.00186)
0 or 1 year *before* birth-year	0.00376**	-0.0632**	-0.0164***	-0.00326	0.00448*	0.00313
	(0.00177)	(0.0312)	(0.00467)	(0.00560)	(0.00237)	(0.00231)
1 or 2 years *after* birth-year	0.00190	-0.0769*	-0.0167***	0.000574	0.00168	-8.22e-05
	(0.00195)	(0.0400)	(0.00517)	(0.00620)	(0.00223)	(0.00196)
3 or 4 years *after* birth-year	0.00278	-0.0545	-0.0124**	0.00469	0.00231	0.00161
	(0.00192)	(0.0339)	(0.00597)	(0.00623)	(0.00309)	(0.00298)
Severe typhoon post-1965						
2 or 3 years *before* birth-year	-0.0129*	0.0524	0.0225	0.0384**	0.00693	0.00595
	(0.00763)	(0.104)	(0.0138)	(0.0151)	(0.00534)	(0.00466)
0 or 1 year *before* birth-year	-0.00210	-0.141***	-0.0105	-0.0852***	-0.0383***	-0.0299***
	(0.00170)	(0.0517)	(0.00661)	(0.0170)	(0.00732)	(0.00572)
1 or 2 years *after* birth-year	-0.000980	-0.0742	-0.00970	-0.0517***	-0.0252***	-0.0207***
	(0.00165)	(0.0454)	(0.00625)	(0.0105)	(0.00494)	(0.00400)
3 or 4 years *after* birth-year	-0.00137	-0.0249	-0.00788	-0.0224***	-0.00826**	-0.00694***
	(0.00167)	(0.0413)	(0.00540)	(0.00698)	(0.00343)	(0.00268)
Observations	2,290,886	2,255,017	2,255,017	2,093,804	2,093,804	2,093,804
R-squared	0.143	0.188	0.135	0.146	0.045	0.037

Notes: Each column shows the results from a separate regression where the outcome variable is regressed on eight dummy variables indicating whether a severe pre-1965 or post-1965 typhoon passed each municipality-birth-year 2 to 3 years before the birth-year / 0 to 1 year before the birth-year / 1 to 2 years after the birth-year / 3 to 4 years after the birth-year, as well as municipality fixed effects, age-by-island-group fixed effects, region-specific time trends, and a dummy for being male. The regressions are run at the individual level. Data source is the 1990 Census of the Philippines 10% Housing Survey. The sample includes all individuals between the ages of 18 and 43 with non-missing information about the municipality of birth in the 10% housing sample of CPH1990. Robust standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

7.6 Alternative Cut-Off Year – 1968

Table A.10: Effects on Cohort Size - Before and After 1968

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: Ages Cohorts 2 to 43				**Sample: Ages Cohorts 18 to 43**			
	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male
Small typhoon in utero	-0.00829	-0.0102	-0.00899	-0.000757	-0.0109	-0.0122	-0.0116	-0.000641
pre-1968	(0.00849)	(0.0105)	(0.0110)	(0.00286)	(0.00780)	(0.00996)	(0.0103)	(0.00285)
Small typhoon year 1&2	0.00596	0.00462	0.00572	-0.000249	0.00416	0.00332	0.00374	-0.000107
pre-1968	(0.00435)	(0.00527)	(0.00503)	(0.00116)	(0.00375)	(0.00489)	(0.00445)	(0.00120)
Severe typhoon in utero	-0.108***	-0.148***	-0.0791*	-0.0188	-0.0987***	-0.134**	-0.0719	-0.0172
pre-1968	(0.0348)	(0.0517)	(0.0450)	(0.0134)	(0.0349)	(0.0527)	(0.0461)	(0.0136)
Severe typhoon year 1&2	0.000267	-0.00536	-0.00110	4.82e-05	0.00341	0.00199	-9.72e-05	0.00157
pre-1968	(0.0177)	(0.0265)	(0.0179)	(0.00621)	(0.0160)	(0.0251)	(0.0170)	(0.00648)
Small typhoon in utero	-0.00902	-0.00308	-0.0185**	0.00372*	-0.0147	0.00714	-0.0414*	0.0109**
post-1968	(0.00675)	(0.00783)	(0.00839)	(0.00196)	(0.0170)	(0.0195)	(0.0225)	(0.00510)
Small typhoon year 1&2	-0.00287	-0.00862**	0.00343	-0.00300***	-0.0184***	-0.0210**	-0.0134	-0.00173
post-1968	(0.00347)	(0.00382)	(0.00408)	(0.000830)	(0.00689)	(0.00850)	(0.00826)	(0.00212)
Severe typhoon in utero	-0.00759	-0.0245	0.0125	-0.00852	-0.0275	-0.0614	0.00570	-0.0154
post-1968	(0.0368)	(0.0422)	(0.0378)	(0.00779)	(0.0538)	(0.0786)	(0.0479)	(0.0171)
Severe typhoon year 1&2	0.0131	0.00296	0.0251	-0.00527	0.00778	-0.00278	0.0193	-0.00523
post-1968	(0.0149)	(0.0153)	(0.0185)	(0.00366)	(0.0184)	(0.0218)	(0.0270)	(0.00726)
Observations	62,286	62,286	62,286	62,286	38,558	38,558	38,558	38,558
R-squared	0.943	0.910	0.906	0.033	0.939	0.896	0.895	0.045

Notes: All details are the same as for Table 3 except that the typhoon treatment variables here are interacted with a dummy variable indicating whether the typhoon passed through before or after December 1968.

Table A.11: Effects on Education - Before and After 1968

	(1)	(2)	(3)
	Literacy	Years of Education	Completed High Sch.
Small typhoon in utero	0.00105	0.0142	-0.000530
pre-1968	(0.000732)	(0.0152)	(0.00205)
Small typhoon year 1&2	-0.000273	-0.00310	-0.00180
pre-1968	(0.000376)	(0.00840)	(0.00110)
Severe typhoon in utero	0.00353	-0.0165	-0.00906
pre-1968	(0.00388)	(0.0656)	(0.00909)
Severe typhoon year 1&2	0.000772	-0.0649*	-0.0152***
pre-1968	(0.00194)	(0.0380)	(0.00517)
Small typhoon in utero	-0.00198	0.0114	0.00118
post-1968	(0.00143)	(0.0275)	(0.00348)
Small typhoon year 1&2	-0.000317	0.0141	0.000798
post-1968	(0.000749)	(0.0162)	(0.00187)
Severe typhoon in utero	-0.00759*	-0.360***	-0.0215
post-1968	(0.00440)	(0.114)	(0.0144)
Severe typhoon year 1&2	0.000368	-0.0507	-0.00780
post-1968	(0.00200)	(0.0486)	(0.00687)
Observations	2,290,886	2,255,017	2,255,017
R-squared	0.143	0.188	0.135

Notes: All details are the same as for Table 4 except that the typhoon treatment variables here are interacted with a dummy variable indicating whether the typhoon passed through before or after December 1968.

Table A.12: Effects on Occupation - Before and After 1968

	(1) Skilled Occupation	(2) Associate Professional	(3) Professional	(4) Skilled Occupation	(5) Associate Professional	(6) Professional
Small typhoon in utero	0.00565**	0.00165	0.00161	0.00516**	0.00143	0.00131
pre-1965	(0.00270)	(0.00139)	(0.00118)	(0.00254)	(0.00127)	(0.00108)
Small typhoon year 1&2	0.00463***	0.00232***	0.00182**	0.00473***	0.00247***	0.00195***
pre-1965	(0.00148)	(0.000883)	(0.000759)	(0.00137)	(0.000821)	(0.000713)
Severe typhoon in utero	0.00322	0.0103**	0.00891*	0.00209	0.00935**	0.00782*
pre-1965	(0.00981)	(0.00461)	(0.00455)	(0.00985)	(0.00441)	(0.00444)
Severe typhoon year 1&2	0.00246	0.00156	0.000101	0.00448	0.00311*	0.00158
pre-1965	(0.00515)	(0.00199)	(0.00213)	(0.00506)	(0.00186)	(0.00173)
Small typhoon in utero	0.0193**	0.00613*	0.00494*	0.0189**	0.00569*	0.00474*
post-1965	(0.00951)	(0.00349)	(0.00274)	(0.00916)	(0.00313)	(0.00243)
Small typhoon year 1&2	0.00736	0.00304	0.00284	0.00687	0.00268	0.00258
post-1965	(0.00454)	(0.00240)	(0.00203)	(0.00428)	(0.00222)	(0.00189)
Severe typhoon in utero	-0.176***	-0.0755***	-0.0571***	-0.162***	-0.0662***	-0.0496***
post-1965	(0.0451)	(0.0184)	(0.0134)	(0.0428)	(0.0165)	(0.0118)
Severe typhoon year 1&2	-0.0572***	-0.0280***	-0.0230***	-0.0568***	-0.0277***	-0.0227***
post-1965	(0.0118)	(0.00527)	(0.00403)	(0.0116)	(0.00513)	(0.00398)
Years of education	No	No	No	0.0381***	0.0252***	0.0215***
				(0.000665)	(0.000711)	(0.000615)
Observations	2,093,804	2,093,804	2,093,804	2,069,113	2,069,113	2,069,113
R-squared	0.147	0.045	0.037	0.232	0.146	0.125

Notes: Details are the same as in Table A.11.

7.7 Alternative Distance to the Eye of the Storm

This subsection presents results with additional treatment variables assigned to municipalities between 30 and 60 kilometers from the eye of the storm. Other than the eight treatment variables in the main analysis, which are assigned to municipalities and individuals within 30 kilometers of the storm path, we add eight more treatment variables, constructed in the same fashion as in the main analysis but assigned to municipalities and individuals that were between 30 and 60 kilometers from the eye of the storm. Specifically, a municipality is exposed to a typhoon 30 to 60 kilometers away if the distance between the centroid of the municipality and the typhoon path is larger than 30 km but no larger than 60 km.

As such, we estimate the effects of typhoon exposure on municipalities and individuals within 30 kilometers of the typhoon path as well as municipalities and individuals between 30 and 60 kilometers from the typhoon path. The "control group" here consists, thus, of municipalities and individuals beyond of the 60-kilometer radius of the storm path.

Tables A.13 presents the results for cohort size and the fraction of males. Table A.14 presents the results for educational attainment and occupational skill levels.

Table A.13: Effects on Cohort Size - Varying Distance to Storm - Never-Zero Municipalities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: Ages Cohorts 2 to 43				**Sample: Ages Cohorts 18 to 43**			
	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male
Small typhoon in utero	-0.00966	-0.0101	-0.0126	1.02e-05	-0.0138	-0.0137	-0.0167	0.000115
within 30km, pre-1965	(0.0103)	(0.0125)	(0.0128)	(0.00310)	(0.00924)	(0.0118)	(0.0116)	(0.00309)
Small typhoon year 1&2	0.00989*	0.00720	0.0110*	-0.000895	0.00718	0.00497	0.00824	-0.000786
within 30km, pre-1965	(0.00535)	(0.00646)	(0.00599)	(0.00132)	(0.00459)	(0.00596)	(0.00520)	(0.00136)
Severe typhoon in utero	-0.103***	-0.151***	-0.0665	-0.0227	-0.0928***	-0.137**	-0.0563	-0.0219
within 30km, pre-1965	(0.0360)	(0.0565)	(0.0461)	(0.0147)	(0.0359)	(0.0572)	(0.0474)	(0.0150)
Severe typhoon year 1&2	0.0105	0.00487	0.00985	0.000155	0.00869	0.00832	0.00554	0.00205
within 30km, pre-1965	(0.0186)	(0.0283)	(0.0184)	(0.00658)	(0.0168)	(0.0272)	(0.0170)	(0.00688)
Small typhoon in utero	-0.0101	-0.00567	-0.0178**	0.00291	-0.0133	-0.000194	-0.0291*	0.00645
within 30km, post-1965	(0.00641)	(0.00748)	(0.00822)	(0.00198)	(0.0122)	(0.0146)	(0.0173)	(0.00440)
Small typhoon year 1&2	-0.00359	-0.00776**	0.000621	-0.00210**	-0.0155***	-0.0154**	-0.0157**	0.000156
within 30km, post-1965	(0.00322)	(0.00367)	(0.00383)	(0.000854)	(0.00507)	(0.00643)	(0.00623)	(0.00169)
Severe typhoon in utero	-0.0241	-0.0371	-0.00810	-0.00658	-0.0545	-0.0913	-0.0182	-0.0167
within 30km, post-1965	(0.0370)	(0.0414)	(0.0391)	(0.00757)	(0.0555)	(0.0681)	(0.0620)	(0.0150)
Severe typhoon year 1&2	0.00424	-0.00724	0.0166	-0.00577	-0.00509	-0.0193	0.00527	-0.00613
within 30km, post-1965	(0.0150)	(0.0155)	(0.0187)	(0.00373)	(0.0212)	(0.0221)	(0.0310)	(0.00720)
Small typhoon in utero	-0.0122	-0.00820	-0.0169	0.00212	-0.0130	-0.00806	-0.0184	0.00247
30 to 60km, pre-1965	(0.00914)	(0.0107)	(0.0125)	(0.00297)	(0.00841)	(0.0101)	(0.0118)	(0.00300)
Small typhoon year 1&2	0.00543	0.00409	0.00516	-0.000250	0.00418	0.00335	0.00328	-8.18e-06
30 to 60km, pre-1965	(0.00495)	(0.00586)	(0.00587)	(0.00130)	(0.00437)	(0.00544)	(0.00529)	(0.00131)
Severe typhoon in utero	-0.0366	-0.0268	-0.0288	-0.00205	-0.0395	-0.0262	-0.0322	-0.00103
30 to 60km, pre-1965	(0.0323)	(0.0425)	(0.0509)	(0.0152)	(0.0328)	(0.0419)	(0.0515)	(0.0151)
Severe typhoon year 1&2	0.0280	0.0182	0.0349	-0.00343	0.0196	0.0135	0.0249	-0.00212
30 to 60km, pre-1965	(0.0175)	(0.0243)	(0.0235)	(0.00730)	(0.0161)	(0.0245)	(0.0217)	(0.00749)
Small typhoon in utero	-0.00599	-0.00575	-0.00797	0.000446	0.00509	0.0118	-0.00412	0.00362
30 to 60km, post-1965	(0.00560)	(0.00662)	(0.00721)	(0.00188)	(0.0102)	(0.0116)	(0.0145)	(0.00377)
Small typhoon year 1&2	-0.000661	0.000923	-0.00174	0.000629	-0.0109**	-0.0113**	-0.0112*	-8.56e-05
30 to 60km, post-1965	(0.00305)	(0.00340)	(0.00374)	(0.000832)	(0.00456)	(0.00571)	(0.00592)	(0.00159)
Severe typhoon in utero	-0.0378*	-0.0538**	-0.0259	-0.00646	-0.0213	-0.0997**	0.0596	-0.0373***
30 to 60km, pre-1965	(0.0204)	(0.0242)	(0.0281)	(0.00747)	(0.0311)	(0.0419)	(0.0457)	(0.0137)
Severe typhoon year 1&2	-0.0188	-0.0236*	-0.0143	-0.00219	-0.00335	-0.00806	-1.49e-05	-0.00183
30 to 60km, post-1965	(0.0122)	(0.0136)	(0.0139)	(0.00276)	(0.0180)	(0.0218)	(0.0239)	(0.00601)
Observations	62,286	62,286	62,286	62,286	38,558	38,558	38,558	38,558
R-squared	0.943	0.910	0.906	0.033	0.939	0.896	0.895	0.045

Notes: Sample restricted to municipalities with at least one male and one female in each cohort under 43. Treatment variables include all those in Table 3 and an additional set indicating exposure to typhoons 30 to 60 km away. A municipality is exposed to a typhoon 30 to 60 km away if the distance between the centroid of the municipality and the typhoon path is between 30km and 60km. All regressions include municipality fixed effects, birth-year by island group fixed effects, and region-specific time trends. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table A.14: Effects on Education and Occupation - Varying Distance to Eye of Storm

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy	Years of Educ.	High Sch.	Skilled Occ.	Associate Prof.	Professional
Small typhoon in utero	0.00150	0.0326*	0.000631	0.00853***	0.00307**	0.00261*
within 30km, pre-1965	(0.000938)	(0.0181)	(0.00244)	(0.00310)	(0.00154)	(0.00134)
Small typhoon year 1&2	0.000176	0.0116	-0.000206	0.00617***	0.00350***	0.00290***
within 30km, pre-1965	(0.000484)	(0.00913)	(0.00125)	(0.00176)	(0.00103)	(0.000887)
Severe typhoon in utero	0.00399	-0.0363	-0.0155*	-0.00398	0.00892*	0.00733
within 30km, pre-1965	(0.00413)	(0.0683)	(0.00899)	(0.0105)	(0.00459)	(0.00447)
Severe typhoon year 1&2	0.000994	-0.0623	-0.0144***	0.000115	0.000891	-0.000914
within 30km, pre-1965	(0.00208)	(0.0389)	(0.00543)	(0.00539)	(0.00209)	(0.00215)
Small typhoon in utero	-0.00156	-0.00949	-0.000548	0.00938	0.00232	0.00215
within 30km, post-1965	(0.00118)	(0.0227)	(0.00289)	(0.00735)	(0.00269)	(0.00207)
Small typhoon year 1&2	-0.000825	-0.00546	-0.00111	0.00426	0.00146	0.00139
within 30km, post-1965	(0.000626)	(0.0132)	(0.00155)	(0.00322)	(0.00165)	(0.00139)
Severe typhoon in utero	-0.0104**	-0.345***	-0.0125	-0.165***	-0.0702***	-0.0544***
within 30km, post-1965	(0.00449)	(0.107)	(0.0138)	(0.0426)	(0.0174)	(0.0133)
Severe typhoon year 1&2	-0.00148	-0.0820*	-0.00978	-0.0539***	-0.0281***	-0.0239***
within 30km, post-1965	(0.00229)	(0.0469)	(0.00680)	(0.0114)	(0.00492)	(0.00397)
Small typhoon in utero	0.00116	0.0182	0.00113	-0.00294	-0.000407	-0.000329
30 to 60km, pre-1965	(0.000932)	(0.0187)	(0.00237)	(0.00243)	(0.00153)	(0.00132)
Small typhoon year 1&2	0.000845*	0.0162*	0.00185*	-0.000594	0.000694	0.00121*
30 to 60km, pre-1965	(0.000492)	(0.00847)	(0.00109)	(0.00119)	(0.000754)	(0.000692)
Severe typhoon in utero	-0.00384	0.0155	-0.00162	-0.00471	-0.000669	0.00184
30 to 60km, pre-1965	(0.00411)	(0.0810)	(0.0104)	(0.00962)	(0.00534)	(0.00487)
Severe typhoon year 1&2	-0.00305	-0.0556	-0.00629	-0.00746	-0.00158	-0.00204
30 to 60km, pre-1965	(0.00260)	(0.0367)	(0.00525)	(0.00476)	(0.00241)	(0.00218)
Small typhoon in utero	-0.00105	0.00577	4.89e-05	-0.00195	0.000532	-0.000226
30 to 60km, post-1965	(0.00112)	(0.0257)	(0.00301)	(0.00733)	(0.00306)	(0.00245)
Small typhoon year 1&2	-0.000888*	0.00889	0.00275	0.000605	0.000133	-0.000444
30 to 60km, post-1965	(0.000533)	(0.0127)	(0.00176)	(0.00280)	(0.00125)	(0.00108)
Severe typhoon in utero	-0.00412	-0.0165	0.0120	-0.0152	-0.00347	-0.00494
30 to 60km, pre-1965	(0.00450)	(0.0792)	(0.0101)	(0.0240)	(0.00780)	(0.00682)
Severe typhoon year 1&2	-0.00156	-0.0540	-0.00317	-0.00861	-0.00708**	-0.00760***
30 to 60km, post-1965	(0.00174)	(0.0408)	(0.00623)	(0.00791)	(0.00332)	(0.00291)
Observations	2,290,886	2,255,017	2,255,017	2,093,804	2,093,804	2,093,804
R-squared	0.143	0.188	0.135	0.146	0.045	0.037

Notes: Sample includes all individuals between the ages of 18 and 43 with non-missing information about the municipality of birth in the 10% housing sample of CPH1990. Definitions of treatment variables are the same as in the main analysis, with an additional set of treatment variables indicating exposure to typhoons "30 to 60 km" away. A municipality is exposed to a typhoon "30 to 60 km" away if the distance between the centroid of the municipality and the typhoon path is larger than 30km but no larger than 60km. All regressions include municipality fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

7.8 Alternative Storm Intensity Measures

In this subsection, we separately estimate the effects of exposure to each category of typhoons on the Saffir-Simpson scale. Instead of using two storm intensity categories (small and severe), we now allow four different storm intensity categories. We continue to combine SS scale 4 and scale 5 storms in one category because only a small number of municipalities were ever exposed to scale 5 storms.

We find evidence for a strong dose-response effect (Table A.15). Before 1965, scale 1 and 2 storms have no significant effect on overall cohort size, while scale 3 and higher storms are associated with smaller cohort size. In general, the magnitude of the adverse effects increases with storm intensity. Moreover, compared to the severe (scale 4 and 5) storms, the effects of scale 3 storms, albeit statistically significant, are much smaller in magnitude. For post-1965 typhoons, we do not find any significant effect on cohort size when we pool all age cohorts of 2- to 43-year-olds together. When we use the restricted sample of 18 to 43-year-olds, we find that category 3 storms are associated with a 7 percent decrease in cohort size. This may be a result of the storms not being "severe enough" for disaster relief funds.

Table A.16 presents the results on education and occupation. Again, in most cases, the magnitude of the adverse effects increases with storm intensity. For pre-1965 typhoons, we do not find any significant reduction in either educational attainment or occupational skill levels among those exposed in early life. For post-1965 typhoons, although we do find statistically significant effects of scale 2 and 3 storms for some outcome variables, the magnitude of the effects is much smaller than the effects of scale 4 and 5 storms.

Table A.15: Effects on Cohort Size with Alternative Storm Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sample: Ages Cohorts 2 to 43				**Sample: Ages Cohorts 18 to 43**			
	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male	ln(Cohort Size)	ln(Male Cohort)	ln(Female Cohort)	Fraction Male
SS Scale 1 in utero	-0.00196	-0.00525	0.00179	-0.00232	-0.00626	-0.00810	-0.00331	-0.00180
pre-1965	(0.0110)	(0.0138)	(0.0134)	(0.00339)	(0.00968)	(0.0128)	(0.0122)	(0.00338)
SS Scale 1 year 1&2	0.00886	0.00142	0.0143**	-0.00302*	0.00649	-0.000291	0.0118*	-0.00283*
pre-1965	(0.00601)	(0.00711)	(0.00706)	(0.00156)	(0.00524)	(0.00660)	(0.00633)	(0.00160)
SS Scale 2 in utero	0.0134	0.0451	-0.0254	0.0159*	0.0117	0.0428	-0.0273	0.0158**
pre-1965	(0.0220)	(0.0288)	(0.0295)	(0.00814)	(0.0208)	(0.0287)	(0.0273)	(0.00802)
SS Scale 2 year 1&2	0.0264**	0.0350**	0.0146	0.00480	0.0253**	0.0335**	0.0133	0.00472
pre-1965	(0.0131)	(0.0159)	(0.0154)	(0.00358)	(0.0109)	(0.0142)	(0.0134)	(0.00358)
SS Scale 3 in utero	-0.0481**	-0.0702**	-0.0445	-0.00608	-0.0527**	-0.0767**	-0.0458	-0.00733
pre-1965	(0.0221)	(0.0294)	(0.0300)	(0.00853)	(0.0216)	(0.0301)	(0.0288)	(0.00865)
SS Scale 3 year 1&2	-0.00941	-0.000155	-0.0134	0.00298	-0.0134	-0.00402	-0.0162	0.00270
pre-1965	(0.0104)	(0.0128)	(0.0134)	(0.00364)	(0.00967)	(0.0125)	(0.0126)	(0.00372)
SS Scale 4 or 5 in utero	-0.0987***	-0.147***	-0.0642	-0.0222	-0.0929***	-0.139**	-0.0573	-0.0219
pre-1965	(0.0352)	(0.0543)	(0.0455)	(0.0142)	(0.0353)	(0.0550)	(0.0468)	(0.0143)
SS Scale 4 or 5 year 1&2	0.0108	0.00712	0.00914	0.000790	0.00895	0.0103	0.00499	0.00257
pre-1965	(0.0186)	(0.0282)	(0.0182)	(0.00643)	(0.0167)	(0.0269)	(0.0167)	(0.00673)
SS Scale 1 in utero	-0.0100	0.00431	-0.0272***	0.00743***	-0.00449	0.0213	-0.0328	0.0123**
post-1965	(0.00738)	(0.00881)	(0.00966)	(0.00251)	(0.0142)	(0.0169)	(0.0209)	(0.00537)
SS Scale 1 year 1&2	-0.00717**	-0.0104**	-0.00390	-0.00158	-0.0137**	-0.0161**	-0.0105	-0.00117
post-1965	(0.00363)	(0.00423)	(0.00429)	(0.000998)	(0.00605)	(0.00777)	(0.00709)	(0.00196)
SS Scale 2 in utero	-0.00359	-0.0231	0.0117	-0.00785**	-0.0244	-0.0499*	-0.00341	-0.0112
post-1965	(0.0152)	(0.0164)	(0.0191)	(0.00387)	(0.0263)	(0.0298)	(0.0352)	(0.00792)
SS Scale 2 year 1&2	0.00739	0.00121	0.0129	-0.00310*	-0.00958	0.00706	-0.0288**	0.00834**
post-1965	(0.00663)	(0.00730)	(0.00792)	(0.00165)	(0.0113)	(0.0137)	(0.0134)	(0.00326)
SS Scale 3 in utero	-0.0223	-0.0386**	-0.00670	-0.00752	-0.0687**	-0.0816*	-0.0542	-0.00604
post-1965	(0.0137)	(0.0175)	(0.0196)	(0.00545)	(0.0313)	(0.0420)	(0.0384)	(0.0112)
SS Scale 3 year 1&2	-0.00208	-0.0126	0.0101	-0.00549**	-0.0309*	-0.0425**	-0.0174	-0.00642
post-1965	(0.00751)	(0.00909)	(0.00904)	(0.00230)	(0.0162)	(0.0214)	(0.0199)	(0.00551)
SS Scale 4 or 5 in utero	-0.0182	-0.0289	-0.00368	-0.00572	-0.0398	-0.0453	-0.0356	-0.00173
post-1965	(0.0360)	(0.0405)	(0.0380)	(0.00755)	(0.0510)	(0.0663)	(0.0544)	(0.0146)
SS Scale 4 or 5 year 1&2	0.00764	-0.00316	0.0194	-0.00545	-0.000385	-0.0117	0.00740	-0.00483
post-1965	(0.0145)	(0.0149)	(0.0182)	(0.00368)	(0.0180)	(0.0203)	(0.0270)	(0.00695)
Observations	62,286	62,286	62,286	62,286	38,558	38,558	38,558	38,558
R-squared	0.943	0.910	0.906	0.034	0.939	0.896	0.895	0.046

Notes: Sample restricted to municipalities that have at least one male and one female in each age cohort under the age of 43. Definitions of treatment variables are the same as in the main analysis except that we now separate storms into four categories – Saffir-Simpson Scale 1, Scale 2, Scale 3, and Scale 4 or 5. All regressions include municipality fixed effects, birth-year by island group fixed effects, and region-specific time trends. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

Table A.16: Effects on Education and Occupation with Alternative Storm Measures

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy	Years of Educ.	High Sch.	Skilled Occ.	Associate Prof.	Professional
SS Scale 1 in utero	0.00186*	0.0486**	0.00205	0.0105***	0.00401**	0.00317**
pre-1965	(0.000989)	(0.0201)	(0.00263)	(0.00335)	(0.00168)	(0.00145)
SS Scale 1 year 1&2	0.000139	0.0149	-0.000665	0.00724***	0.00338***	0.00236**
pre-1965	(0.000541)	(0.0110)	(0.00148)	(0.00195)	(0.00120)	(0.00103)
SS Scale 2 in utero	-4.40e-05	-0.0491	-0.00716	-0.00246	0.00223	0.00424
pre-1965	(0.00189)	(0.0496)	(0.00689)	(0.00664)	(0.00445)	(0.00387)
SS Scale 2 year 1&2	-0.000328	-0.0314	-0.00268	-0.00110	0.00210	0.00196
pre-1965	(0.00118)	(0.0282)	(0.00378)	(0.00316)	(0.00191)	(0.00170)
SS Scale 3 in utero	-0.000696	-0.00484	-0.00194	0.0109	-0.00154	-0.00223
pre-1965	(0.00205)	(0.0383)	(0.00511)	(0.00668)	(0.00311)	(0.00255)
SS Scale 3 year 1&2	-0.000919	0.0123	0.00143	0.00809**	0.00344	0.00368*
pre-1965	(0.00106)	(0.0216)	(0.00288)	(0.00379)	(0.00209)	(0.00191)
SS Scale 4 or 5 in utero	0.00397	-0.0401	-0.0140	-0.00402	0.00824*	0.00691
pre-1965	(0.00397)	(0.0663)	(0.00873)	(0.0100)	(0.00456)	(0.00450)
SS Scale 4 or 5 year 1&2	0.00134	-0.0580	-0.0139***	0.000488	0.000868	-0.000733
pre-1965	(0.00209)	(0.0387)	(0.00529)	(0.00534)	(0.00216)	(0.00224)
SS Scale 1 in utero	9.18e-05	0.00373	-0.00265	0.0259***	0.00873***	0.00703***
post-1965	(0.00162)	(0.0278)	(0.00329)	(0.00877)	(0.00312)	(0.00234)
SS Scale 1 year 1&2	0.000314	-0.00270	-0.00249	0.00680*	0.00333*	0.00331**
post-1965	(0.000701)	(0.0152)	(0.00187)	(0.00374)	(0.00176)	(0.00147)
SS Scale 2 in utero	-0.00481**	-0.0592	0.00291	-0.0391***	-0.0180***	-0.0130***
post-1965	(0.00200)	(0.0491)	(0.00652)	(0.0116)	(0.00530)	(0.00422)
SS Scale 2 year 1&2	-0.00229**	0.0106	0.000131	-1.59e-05	-0.00210	-0.00219
post-1965	(0.00111)	(0.0228)	(0.00285)	(0.00549)	(0.00284)	(0.00232)
SS Scale 3 in utero	-0.00542*	-0.00468	0.00350	0.000437	0.000288	-0.000180
post-1965	(0.00300)	(0.0718)	(0.00816)	(0.0208)	(0.00698)	(0.00535)
SS Scale 3 year 1&2	-0.00281*	-0.0672*	0.000612	0.00603	-0.000626	-0.000681
post-1965	(0.00144)	(0.0374)	(0.00521)	(0.00819)	(0.00378)	(0.00314)
SS Scale 4 or 5 in utero	-0.00737*	-0.308***	-0.0184	-0.152***	-0.0659***	-0.0499***
post-1965	(0.00398)	(0.106)	(0.0143)	(0.0395)	(0.0166)	(0.0126)
SS Scale 4 or 5 year 1&2	-0.000218	-0.0688	-0.00989	-0.0491***	-0.0243***	-0.0197***
post-1965	(0.00186)	(0.0443)	(0.00651)	(0.0111)	(0.00495)	(0.00391)
Observations	2,290,886	2,255,017	2,255,017	2,093,804	2,093,804	2,093,804
R-squared	0.143	0.188	0.135	0.147	0.045	0.037

Notes: Sample includes all individuals between the ages of 18 and 43 with non-missing information about the municipality of birth in the 10% housing sample of CPH1990. Definitions of treatment variables are the same as in the main analysis except that we now separate storms into four categories – Saffir-Simpson Scale 1, Scale 2, Scale 3, and Scale 4 or 5. All regressions include municipality fixed effects, birth-year by island group fixed effects, region-specific time trends, and a dummy for being male. Standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

7.9 Effects of Severe Typhoons by Lustrum[35]

Table A.17: Effects of Severe Typhoons by Lustrum - Cohort Size

	(1) ln(Cohort Size)	(2) ln(Male)	(3) ln(Female)	(4) Fraction Male
In utero exposure to				
severe typhoons 1955 or earlier	-0.0781	-0.188*	0.0320	-0.0504**
	(0.0653)	(0.108)	(0.0553)	(0.0227)
severe typhoons 1956 -1960	-0.0936	-0.113	-0.151	-0.00716
	(0.0837)	(0.0905)	(0.156)	(0.0306)
severe typhoons 1961 -1965	-0.0996**	-0.106*	-0.1000	-0.00169
	(0.0479)	(0.0569)	(0.0678)	(0.0176)
severe typhoons 1966 or later	-0.0127	-0.0302	0.00929	-0.00918
	(0.0373)	(0.0418)	(0.0395)	(0.00771)
Observations	62,286	62,286	62,286	62,286
R-squared	0.943	0.910	0.906	0.033

Notes: Each column shows the results from a separate regression where the outcome variable is regressed on the four typhoon exposure variables shown here as well as well as municipality fixed effects, age-by-island-group fixed effects, and region-specific time trends. The four typhoon exposure variables shown here measure the expected number of severe typhoons in each of the four time periods that passed within 30 km of the cohort's municipality of birth when the cohort was in utero. The regressions are run at the municipality-age-cohort level. Data source is the 1990 Census of the Philippines 10% Housing Survey. Sample includes cohorts between the ages of 2 and 43 in 1990 and municipalities with at least one male and one female in each age cohort for all ages under 43 in the 1990 Census 10% Housing Survey. Robust standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

[35]For simplicity, the regressions presented here omit early-life exposure to low-intensity typhoons as well as exposures in the first two years of life. Including the full set of treatment variables affect the key coefficients little.

Table A.18: Effects of Severe Typhoons by Lustrum - Education and Occupation

	(1) Literacy	(2) Years of Education	(3) Completed High Sch.	(4) Skilled Occupation	(5) Associate Professional	(6) Professional
In utero exposure to						
severe typhoons 1955 or earlier	0.00794	0.0172	-0.0208**	0.0156	0.0234**	0.0231**
	(0.00742)	(0.100)	(0.00855)	(0.0148)	(0.0107)	(0.0102)
severe typhoons 1956 -1960	0.00971	-0.0173	-0.0118	-0.0690**	-0.00329	0.00137
	(0.00786)	(0.216)	(0.0254)	(0.0280)	(0.00648)	(0.00754)
severe typhoons 1961 -1965	0.00238	-0.103	-0.0216	-0.00167	0.00507	0.000684
	(0.00584)	(0.0864)	(0.0132)	(0.0152)	(0.00465)	(0.00412)
severe typhoons 1966 or later	-0.00653*	-0.359***	-0.0195	-0.191***	-0.0830***	-0.0638***
	(0.00381)	(0.112)	(0.0145)	(0.0439)	(0.0188)	(0.0144)
Observations	2,290,886	2,255,017	2,255,017	2,093,804	2,093,804	2,093,804
R-squared	0.143	0.188	0.135	0.146	0.045	0.037

Notes: Each column shows the results from a separate regression where the outcome variable is regressed on the four typhoon exposure variables shown here as well as well as municipality fixed effects, age-by-island-group fixed effects, and region-specific time trends. The four typhoon exposure variables shown here measure the expected number of severe typhoons in each of the four time periods that passed within 30 km of the respondent's municipality of birth when the respondent was in utero. The regressions are run at the individual level. Data source is the 1990 Census of the Philippines 10% Housing Survey. Sample includes all individuals between the ages of 18 and 43 with non-missing information about the municipality of birth in the 10% housing sample of CPH1990. Robust standard errors are clustered two-way at both the municipality level and the province-by-birth-year level.

8 Data Appendix

8.1 Typhoon Exposure Measures

Exposure variable at the storm-municipality-birth year level

We determine the set of affected municipalities for each storm based on the distance between the centroid of the municipality and the typhoon path. To do this, we first generate best-fit lines through the six-hourly typhoon observations to identify the storm path. We then use ArcGIS to calculate the distance between the centroid of the municipality and the storm path for each municipality-storm pair. If the distance is within 30 kilometers, we treat the municipality as affected by the storm on the day that the typhoon was recorded by the nearest observation point.

Second, we determine storm intensity for each affected municipality. We categorize storm

intensity on the basis of its minimum central pressure (MCP). Storm intensity varies as the storm moves, and we capture this using the six-hourly readings. We use the distance-weighted average of the MCP readings at the two nearest observation points as the MCP measure for the municipality. We categorize storm intensity by MCP as follows:

Saffir-Simpson Scale	Exposure	Minimum central pressure
1	Small	980-1000
2	Small	965-980
3	Small	945-965
4	Severe	920-945
5	Severe	$<$920

Third, for each storm-municipality-birthyear triple, we determine the probability that individuals born in the municipality during the birth-year would be affected by the storm either in utero or during the first two years after birth. We resort to this measure because month of birth was not recorded in CPH 1990. CPH 1990 records each respondent's age as his or her age on May 1, 1990. Most typhoons take place between July and November of each year. The probability of exposure measure allows us to fully exploit variation in the timing of the typhoon.

To compute the probability of exposure, we assume no seasonality of conception in the Philippines. In other words, we assume that an individual born in a given birth-year is equally likely to be born on any day during the year if there were no exogenous shocks to fetal mortality. We formally test this assumption in the next subsection and find that this is indeed the case in the Philippines for individuals born between 1947 and 1972. The expected number of typhoons that a respondent is exposed to depends on his or her municipality of birth and age as of May 1, 1990. We assume 40 weeks of gestation (280 days).[36]

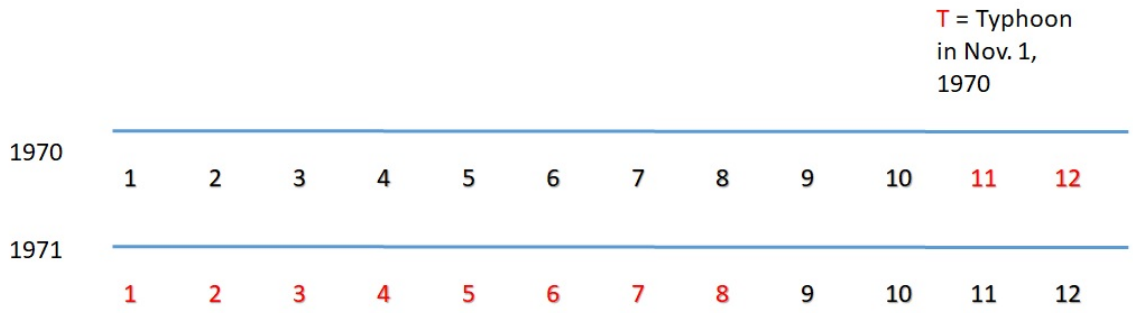
For any given recorded *age*, the respondent could be born either between May 2 and December 31 of the year given by $1990 - age - 1$ or between January 1 and May 1 of the year given by $1990 - age$. For example, an individual who is 19 years old on May 1, 1990

[36]For robustness, we also use 36 weeks and 38 weeks. The results are qualitatively similar.

may be born anywhere between May 2, 1970 and May 1, 1971.

If there were a typhoon on month-date T, year Y, this typhoon would affect the in utero period of all individuals born up to 280 days after month-date T, year Y. For example, consider figure B.1 below. Consider a typhoon that takes place on November 1, 1970 ($T =$ November 1, $Y =$1970).

Figure B.1: Probability of exposure



This typhoon would affect the in utero period of all individuals born between November 1, 1970, and August 8, 1971. The probability of in utero exposure to this typhoon is the fraction of the cohort's potential birth dates that lie between November 1, 1970, and August 8, 1971. This typhoon may have affected the in utero period of both the 18-year-old cohort and the 19-year-old cohort. However, because the "November 1, 1970, to August 8, 1971" window has a larger overlap with the potential birth dates of the 19-year-old cohort, the typhoon has a larger probability of affecting the in utero period of the 19-year-old cohort than the 18-year-old cohort.

Similarly, this typhoon could affect the post-utero period, which we define as the first two years after birth, of all individuals born in the two years before November 1, 1970. It may have affected the post-utero period of the 19-year-old, 20-year-old, and 21-year-old cohorts. For the 20-year-old cohort, the probability of post-utero exposure to this typhoon is one because their birth dates must lie within the two-year window before November 1, 1970. For

the 19-year-old and 21-year-old cohort, the probability of post-utero exposure is then the fraction of the cohort's potential birth dates that lie in the two years before November 1, 1970.

Depending on each person's age, their probability of exposure to a typhoon that occurs on this date is calculated as the following (for a typhoon on November 1, 1970):

$$\text{P(Affected postutero|Age 21 on May 1, 1990)}=\frac{No.\,of\,days\,between\,November\,1,\,1968\,and\,May\,1,\,1969}{365}$$

$$\text{P(Affected postutero|Age 20 on May 1, 1990)}=1$$

$$\text{P(Affected postutero|Age 19 on May 1, 1990)}=\frac{No.\,of\,days\,between\,May\,2,\,1970\,and\,November\,1,\,1970}{365}$$

$$\text{P(Affected inutero|Age 19 on May 1, 1990)}=\frac{No.\,of\,days\,between\,November\,1,\,1970\,and\,May\,1,\,1971}{365}$$

$$\text{P(Affected inutero|Age 18 on May 1, 1990)}=\frac{No.\,of\,days\,between\,May\,2,\,1971\,and\,August\,8,\,1971}{365}$$

We also note that not all typhoons could affect the in utero period of both cohorts. Typhoons in May and June, for example, would affect only one cohort – the cohort born in the same year, because the 280-day window would not expand beyond May 1 of the following year. For example, for a typhoon on June 1, 1970:

$$\text{P(Affected postutero|Age 21 on May 1, 1990)}=\frac{No.\,of\,days\,between\,June\,1,\,1968\,and\,May\,1,\,1969}{365}$$

$$\text{P(Affected postutero}|\text{Age 20 on May 1, 1990)}=1$$

$$\text{P(Affected postutero}|\text{Age 19 on May 1, 1990)}=\frac{No.\,of\,days\,between\,May\,2,\,1970\,and\,June\,1,\,1970}{365}$$

$$\text{P(Affected inutero}|\text{Age 19 on May 1, 1990)}=\frac{280}{365}$$

$$\text{P(Affected inutero}|\text{Age 18 on May 1, 1990)}=0$$

We generate exposure probabilities in a similar manner for all storm-municipality-birthyear triples. We then calculate the expected number of typhoons (by intensity) that each municipality-birthyear pair is exposed to by adding up the probabilities that the municipality-birthyear is exposed to each typhoon in the database.

Seasonality of birth We formally test the seasonality assumption using the 1993 Demographic and Health Survey (1993 DHS), which records respondents' month of birth. We follow the methods in He and Earn (2007) and Dorélien (2016). Using the 1993 DHS, we recover the distribution of month of birth for women aged 18 to 43 in 1990 (i.e. born between 1947 and 1972). As presented in Table B.1, the distribution is indeed close to uniform. To formally test the assumption, we begin by aggregating the number of births in each month of the year, using the 1993 DHS sample weight as the inflation factor. We then convert the month and year of birth data to monthly amplitude, which is defined as the percent deviation from the annual monthly mean as follows:

$$A_{my} = \frac{C_{my}X_{my} - \bar{X}_y}{\bar{X}_y}$$

where

$$\bar{X}_y = \frac{1}{12} \sum X_{my}$$

is the average number of births in a month in year y and $C_{my} = \frac{days\,in\,y/12}{days\,in\,m\,of\,year\,y}$ is the scaling factor to correct for the number of days in each month of the year. We then regress monthly amplitude on birth month indicators to formally test the assumption of equal proportion of births across the 12 months of the year. The regressions (columns 2, 4, and 6) confirm that seasonality is low in the sample: the R-squared from each regression is low and we fail to reject the null hypothesis of equal deviations across the 12 months of the year (with p-values ranging from 0.292 to 0.762).

Table B.1: Seasonality of birth for Women Aged 18 to 43 in 1990

	(1)	(2)	(3)	(4)	(5)	(6)
	Birth year 1947-1972		Birth year 1947-1965		Birth year 1966-1972	
	%	Difference	%	Difference	%	Difference
January	8.71		8.64		8.86	
February	7.55	0.0121*	7.50	0.0141	7.65	0.00669
		(0.00727)		(0.00891)		(0.00851)
March	8.56	0.00491	8.34	0.00724	9.03	-0.00143
		(0.00698)		(0.00868)		(0.00594)
April	8.11	0.00189	8.25	0.00162	7.82	0.00262
		(0.00575)		(0.00712)		(0.00639)
May	8.68	0.000879	8.47	0.00167	9.13	-0.00127
		(0.00683)		(0.00853)		(0.00747)
June	7.82	0.0100	7.50	0.0118	8.50	0.00531
		(0.00702)		(0.00882)		(0.00605)
July	7.71	0.00708	7.82	0.00885	7.46	0.00229
		(0.00879)		(0.0114)		(0.00706)
August	8.33	0.00134	8.46	0.00277	8.05	-0.00255
		(0.00705)		(0.00889)		(0.00628)
September	8.52	-0.00192	8.63	-0.00333	8.28	0.00191
		(0.00548)		(0.00658)		(0.00813)
October	8.62	0.00300	8.72	0.00324	8.42	0.00235
		(0.00823)		(0.0107)		(0.00752)
November	8.68	0.00150	8.88	-4.25e-05	8.28	0.00568
		(0.00564)		(0.00683)		(0.00825)
December	8.71	-0.000878	8.79	-0.00145	8.54	0.000677
		(0.00542)		(0.00663)		(0.00646)
R-squared		0.020		0.030		0.059
F-statistic		0.86		0.68		1.22
p-value		0.762		0.753		0.292

Notes: Authors' calculations using the Philippines Demographic and Health Survey 1993. Regressions are run at the month of birth by birth year level. Difference (cols. 2, 4, 6) captures monthly amplitude, defined as the percent deviation from the annual monthly mean, relative to January. Regressions at the birth month-birth year level. Sample restricted to women born between 1947 and 1972 (aged 18 to 43 in 1990). Sample weight included, robust standard error shown in parentheses.

Measurement errors due to migration before age 2 One concern with our treatment variable is measurement error in the post-utero exposure variables due to migration in the first two years of life. Migration introduces measurement error in our post-utero exposure

variables since we assign the probability of exposure in the first two years of life based on individuals' municipality of birth but we are unable to determine individuals' municipality of residence in their first or second year of life. We first note that, given the high probability of typhoon exposure in the Philippines, it is unlikely that individuals will move to avoid typhoon exposure. Hence, neither the probability nor the direction of migration should be correlated with post-utero typhoon exposure and the measurement error introduced by migration could only bias our estimates downwards. We further note that despite the high rates of migration in the Philippines, the probability of moving in the first two years of life is low. We do not have the complete migration history of all adults between 18 and 43. We find that less than 3% reported migrating to their place of residence in 1985 before the age of 2 and when we interact household SES and typhoon exposure, we find no evidence of selective migration by household SES. Using CPH 1990, we restrict the sample to households with children under 5 and find that about 5% of households migrated in the previous five years. We also use the 1993 Philippines Demographic and Health Survey (1993 DHS) to explore early life migration for women aged 18 to 43 in 1990 (i.e. born between 1947 and 1972). More than 50 percent of women born between 1947 and 1972 in the sample moved at least once in their lives. The 1993 DHS does not include a complete migration history, but we restrict the sample to women who moved to their place of residence in 1993 before the age of 15, and find that 6% of the sample moved to their current place of residence before the age of 2. We repeat the exercise using children under 5 in the 1993 DHS and also find that 5% of children under 5 in the survey moved before the age of 2.

8.2 Outcomes and Covariates

Cohort size We measure cohort size at the birth-municipality birth-year level. The variable is defined as the number of individuals born in municipality m in year t who survived until May 1, 1990. To order to capture cohort size at the municipality of birth (rather than the municipality of residence in 1990), we continue to use the 10% sample of CPH 1990 for

which we observe each individual's municipality of birth. The cohort size of municipality m in year t is calculated by adding up the person weights of all individuals who were born in municipality m in year t.[37] We similarly define the male cohort and the female cohort by restricting the sample to males and females respectively.

Education The 1990 Census categorizes education into pre-school, elementary 1 to 6, high school 1 to 4, and post-secondary education. We count pre-school completion as having 1 year of education. Elementary school completion is equivalent to having 7 years of education and high school completion to 11 years of education. Most students in the Philippines begin formal elementary school at age 6; assuming no grade repetition, respondents should be 16 when they complete high school. We use the Census categories for post-secondary education to convert post-secondary achievement to years of education.

Occupational skill level We categorize occupations into professional, associate professional, and skilled occupations based on the the 1990 Census categories. CPH1990 records occupations using the 4-digit Philippine Standard Occupational Classification (PSOC) system, which is patterned after the International Standard Classification of Occupations (ISCO) released by the International Labour Organization. The PSOC categorizes occupations by both the set of tasks carried out and the skills involved in the job. Each major group in the PSOC can be mapped to a ISCO skill level, which we use to categorize occupations into professional, associate professional, and skilled.

The Census questionnaire asks one's occupation in the past year as well as the past week. Whenever occupation from the past year is available, we use it as one's occupation; when occupation from the past year is not reported, we use occupation from the past week.

We then map occupations into professional, associate professional, and skilled occupations according to the first digit of the 4-digit occupation code. Professional occupations

[37] As a robustness check, we also perform all our analyses using a simple count of individuals as cohort size, rather than adding up the weights of all individuals. All results are very similar.

include two major groups: officials of the government and special interest organizations, corporate executives, managers, managing proprietors and supervisors (PSOC group 1) and professionals (PSOC group 2). Associate professionals include, besides the two groups included in professionals, another major group which requires skills at the third ISCO skill level: technicians and associate professionals (PSOC group 3). Skilled occupations add five other major groups that require skills at the second ISCO skill level: clerks (PSOC group 4); service workers and shop and market sales workers (PSOC group 5); farmers, forestry workers, and fishermen (PSOC group 6); trades and related workers (PSOC group 7); plant and machine operators and assemblers (PSOC group 8). Excluded from the skilled occupations are laborers and unskilled workers (PSOC group 6) which requires skills at the first ISCO skill level. Also excluded from skilled occupations are non-gainful occupations, which includes volunteers, housekeepers, students, pensioners and other retirees, disabled, and other non-gainful or no reported activity.

Using the first digit of PSOC code allows us to capture the required level of skill for one's usual job. For example, in agriculture, the PSOC separates "skilled rice farm worker" from "farm laborer" and we code the two occupations under different skill levels.

9 A Conceptual Framework for Culling and Scarring

We present a modified version of the "fetal origins" model in Almond (2006) to illustrate how in utero exposure to typhoons and changes in post-disaster relief policies could interact to create the culling and scarring effects we have documented in this paper.

Following Almond (2006), we allow a severe typhoon to have two potential effects on fetal health: (1) a negative shift in the unobserved distribution of fetal health (scarring) and (2) a reduction in the survival odds conditional on health. Both effects contribute to culling. The latter may result from deteriorating medical care and a shortage of clean water and food in the ensuing weeks after a severe typhoon. The fetal origins hypothesis asserts that the first effect, the deterioration in the distribution of fetal health, would persist into adulthood and result in poorer educational and labor market outcomes.

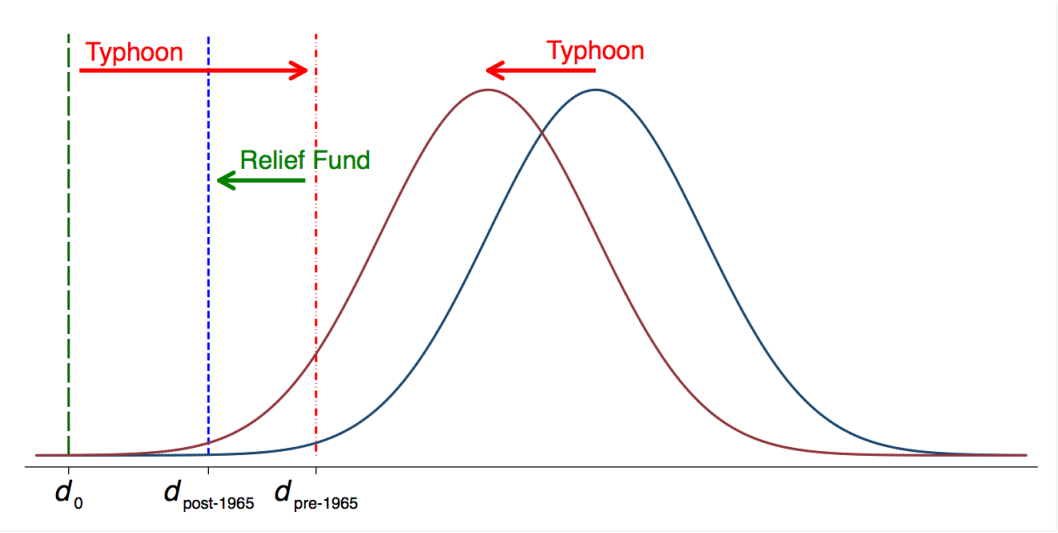
We modify the model to incorporate changes in post-disaster relief policies. We assume that Marcos' post-disaster relief policies alleviated the second effect of typhoons but did not have much impact on the first effect. That is, by providing food, water, and medical services *after* each typhoon incident, Marcos' post-disaster relief policies would reduce the negative effects of typhoons on survival odds conditional on health.

This is illustrated graphically in Figure D.1. Let h denote the underlying fetal health. In the absence of any natural disasters, $F_0(\cdot)$ is the cumulative distribution function of fetal health. There is a natural fetal mortality rate associated with health threshold d_o such that those whose $h < d_o$ would not survive to adulthood. A severe typhoon shifts the distribution of fetal health to the left, to $F_1(\cdot)$. In the absence of any disaster relief measures, a severe typhoon also reduces the survival odds such that those with $h < d_{pre}$ would not survive to adulthood ($d_o < d_{pre}$). If d_{pre} is sufficiently high, $E(h \mid h \geq d_{pre}, F_1(\cdot))$ may be very close to $E(h \mid h \geq d_0, F_0(\cdot))$. In other words, if mortality rate is sufficiently high after the disaster, we may not observe long-term scarring.

Marcos' post-disaster relief policies shift the survival threshold to the left, to d_{post},where

$d_0 \leq d_{post} < d_{pre}$. If the post-disaster relief policies are sufficiently effective to bring d_{post} close to d_0, then the difference between $E(h \mid h \geq d_{post}, F_1(\cdot))$ and $E(h \mid h \geq d_0, F_0(\cdot))$ would better reflect the scarring effects of the disaster. In other words, some of those who would have died in the absence of the disaster relief policy survived as a result of the policy. The change in the composition of respondents who survived pre- and post-1965 results in changes in the observed long-term scarring effects.

Figure D.1: Conceptual framework for culling and scarring



References

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 us population. *Journal of Political Economy* 114(4), 672–712.

Dorélien, A. (2016). Birth seasonality in Sub-Saharan Africa. *Demographic Research* 34, 761–796.

He, D. and D. J. Earn (2007). Epidemiological effects of seasonal oscillations in birth rates. *Theoretical Population Biology* 72(2), 274–291.