California Center for Population Research
University of California - Los Angeles

# Air Quality and Early-Life Mortality During Indonesia's Massive Wildfires in 1997

**Seema Jayachandran**

**CCPR-023-05**

**August 2005**

*California Center for Population Research*
*On-Line Working Paper Series*

# Air Quality and Early-Life Mortality During Indonesia's Massive Wildfires in 1997*

Seema Jayachandran†

August 2005

**Abstract**

Air quality was extremely poor in Indonesia in late 1997 due to smoke from massive wildfires. This paper examines the impact this episode of air pollution (particulate matter) had on infant and fetal mortality. Deaths are inferred from "missing children" in the 2000 Indonesian Census, analyzing subdistrict-year-month birth cohorts and exploiting the sharp timing and spatial patterns of the pollution. Exposure to pollution during the last trimester in utero is found to have a large effect on survival. The fire-induced pollution caused a 1.0% decrease in cohort size, or over 16,400 missing children across Indonesia for the five-month period of high exposure. In addition, pollution has much larger mortality effects in poorer areas. The results suggest that environmental damage that occurs alongside economic development has large and regressive health costs.

*I am grateful to Janet Currie, Erica Field, Dan Kammen, Michael Kremer, Doug McKee, Ben Olken, Sarah Reber, Duncan Thomas, and seminar participants at Berkeley, Columbia, UCLA, and the April 2005 BREAD conference for helpful comments. I also thank Kok-Hoe Chan and Daniel Chen for data and Nadeem Karmali and Rika Christanto for excellent research assistance.

†Robert Wood Johnson Foundation Scholar in Health Policy Research, UC Berkeley; Assistant Professor, Dept. of Economics, UCLA. e-mail: jaya@ucla.edu.

# 1 Introduction

Between September and November 1997, forest fires raged through large parts of Indonesia, destroying over 12 million acres. Most of the fires, which were concentrated on the islands of Sumatra and Borneo (Kalimantan), were started intentionally by logging companies and palm oil producers clearing land to plant new crop.[1] Because of the dry, windy conditions caused by El Niño, the fires burned out of control and spread rapidly. In November, rains finally doused the fires.

While the fires were burning, much of Indonesia was blanketed in smoke. This paper examines infant and fetal mortality caused by the episode of poor air quality (specifically, high levels of particulate matter). Daily satellite measurements of airborne smoke at locations across Indonesia provide information on the spatial and temporal patterns of the pollution. The outcome, infant and fetal mortality, is inferred from "missing children" in the 2000 Census, overcoming the lack of mortality records for Indonesia and the small samples in surveys on infant mortality.

The paper finds that higher levels of pollution cause a substantial decline in the size of the surviving cohort, and that exposure to pollution during the last trimester in utero is the most damaging. The fire-induced increase in air pollution is associated with a 1.0% decrease in cohort size, averaged across Indonesia for the five-month period of high exposure. Indonesia's under–2 mortality rate during this period was 6%; assuming the effect of pollution was mainly on infant deaths (rather than fetal deaths), this represents a 17% increase in under-2 mortality.

The results imply that over 16,400 infant and fetal deaths are attributable to the fires. Most cost estimates of the fires have focused on destroyed timber, reduced worker productivity, lost tourism, and the like and are in the range of $2 to 3 billion (Tacconi 2003). The health costs of the fires are likely much larger: Assuming a value of a statistical life of $1

[1]The Indonesian Minister of Forestry estimated that "[commercial] plantations caused some 80% of the forest fires," and that small farmers caused the remainder (*Straits Times*, September 3, 1997). Rabindran (2001), using satellite data on land use, finds that the 1997 incidence of fires on plantations was higher than the "natural" level (based on a benchmark from conservation areas), but the incidence of fires on small farms was at its natural level.

million, the infant mortality costs alone were over $16 billion. These costs very likely overwhelm the benefits to firms from setting fires; the annual revenue from Indonesia's timber and palm oil industries during this period was less than $7 billion.

The paper also finds a striking difference in the mortality effects of pollution between richer and poorer places. Pollution has twice the effect on mortality in districts with consumption below the sample median compared to those above the median. There are a number of possible explanations for this finding. Individuals in poorer areas could be more susceptible to pollution because of lower baseline health, more limited options for avoiding the pollution, or less access to medical care. Another possibility is that people exposed to indoor air pollution on a daily basis suffered more acute health effects from the wildfires because they received a double dose of pollution. Consistent with this view, the estimated effects are larger in areas where more people cook with wood-burning stoves. In addition, pollution causes more mortality in areas with fewer doctors and medical facilities. Mother's education also seems to play a role. While these correlations do not pin down a causal relationship, they provide some suggestive evidence on why the poor are especially vulnerable to the adverse health effects of pollution.

This paper contributes to the literature by providing evidence on the causal effect of pollution on health in a developing country. Rampant wildfires are frequent in Indonesia and in many other countries in Southeast Asia and Latin America where fire is used to clear land, but the type of air pollution studied here has even broader applicability in poor countries. In particular, pollution from the wildfires is comparable to that from wood-burning stoves which are widely used in developing countries and produce a similar level and mix of pollutants. While indoor air pollution has recently become a focal health issue for agencies such as the World Health Organization, there is still a great deal of uncertainty regarding the magnitude of the health impact from biomass fuel. The estimates in this paper indicate that reductions in indoor stove pollution would save many lives.

The paper sheds light on not just the overall magnitude but also the detailed nature of health damage from pollution. Two contributions stand out. First, the results provide

evidence on when in early life exposure to pollution matters most. In utero exposure is found to be especially important. This suggests that targeting pregnant women should be a priority of public health efforts concerning air pollution. The abruptness of Indonesia's 1997 pollution event makes this study uniquely well suited to studying exposure to pollution at different stages of life. Second, the paper explores differential effects by poverty level, revealing that the adverse health effects of pollution are not only large, but also very regressive. This finding has important welfare and policy implications.

In addition, the episode of air pollution studied in this paper exemplifies two broader phenomena related to the environment in developing countries. First, environmental damage that is occurring alongside industrialization appears to have large health costs. As Indonesia has liberalized its trade and expanded exports of timber and palm oil, fires set by commercial interests and, consequently, outbreaks of widespread pollution have become more prevalent. Second, environmental degradation and its health effects are one of the many consequences of weak governance. The illegal logging and land clearance that contributed to the fires were made possible by Indonesia's lax enforcement of environmental regulations.

The remainder of the paper is organized as follows. Section 2 provides background on the link between pollution and health and on the Indonesian fires. Section 3 describes the data and empirical strategy. Section 4 presents the results, and section 5 concludes.

# 2 Background

## 2.1 Link between air pollution and infant mortality

### Related literature

Recent work on air quality and infant mortality includes that by Chay and Greenstone (2003b) who use geographic variation across the United States in the extent to which the 1980–81 recession lowered pollution. They find that better air quality reduced infant deaths. Chay and Greenstone (2003a) find that air pollution abatement after passage of the Clean

Air Act of 1970 also led to a decline in infant deaths.[2] Currie and Neidell (2005) use within-zipcode variation in California over the 1990's and find that exposure to carbon monoxide and other air pollutants during the month of birth is associated with infant mortality.[3]

In addition, there have been studies on the adult health effects of Indonesia's 1997 fires. Emmanuel (2000) finds no increase in mortality but an increase in respiratory-related hospitalizations in nearby Singapore. Sastry (2002) finds increased mortality for older adults on the day after a high-pollution day in Kuala Lampur and Kuching, Malaysia. Frankenberg, McKee, and Thomas (2004) compare adult health outcomes in 1993 and 1997 for areas in Indonesia with high versus low exposure to the 1997 smoke. They find that pollution reduced people's ability to perform strenuous tasks and other measures of health. The data set used by Frankenberg, McKee, and Thomas (2004) covers only 321 of the nearly 4000 subdistricts in Indonesia, however. Only one of Kalimantan's four provinces is in their sample. Thus, one advantage of this paper is its broader geographic coverage, which allows one to explore heterogeneous effects across households and nonlinearities in the health impact of pollution, for example. In addition, the level of pollutant exposure in Kalimantan most closely approximates exposure from wood–burning stoves, so the comparison offered in this paper provides better estimates of the health impact of indoor air pollution. Another important advance over previous work in Indonesia is that the identification strategy exploits both the sharp timing and regional variation of the pollution.

### Physiological effects of pollution

Smoke from burning wood and vegetation, or biomass smoke, consists of very fine particles (organic compounds and elemental carbon) suspended in gas. Fine particles less than 10 microns ($\mu$m) and especially less than 2.5 $\mu$m in diameter are considered the most harmful to health because they are small enough to be inhaled and transported deep into the lungs.

[2]Other natural experiments that have been used to measure health effects of air pollution include the temporary closure of a steel mill in Utah during a 1986–7 labor dispute and the reduction in traffic during the 1996 Olympics in Atlanta (Pope et al. 1992, Friedman et al. 2001).

[3]For research on pollution and infant mortality outside the U.S., see for example Bobak and Leon (1992) on the Czech Republic, Loomis et al. (1999) on Mexico, and Her Majesty's Public Health Service (1954) on the 1952 London "killer fog" (smog) episode.

For biomass smoke, the modal size of particles is between 0.2 and 0.4 $\mu$m, and 80 to 95% of particles are smaller than 2.5 $\mu$m (Hueglin et al. 1997).

There are several possible pathways through which prenatal and postnatal exposure to air pollution could affect fetal or infant health. Postnatal exposure can contribute to acute respiratory infection, a leading cause of infant mortality. In utero exposure is hypothesized to affect fetal development, first, because pollution inhaled by the mother and absorbed into her bloodstream interferes with her health which in turn disrupts fetal nutrition and oxygen flow, and, second, because toxicants cross the placenta. Several studies find an association between air pollution and fetal growth retardation or shorter gestation period, both of which are associated with low birthweight (Dejmek et al. 1999, Wang et al. 1997, Berkowitz et al. 2003).

There is also evidence on the biological mechanisms behind these pregnancy outcomes. The main toxicant in most particulate matter including biomass smoke is polycyclic aromatic hydrocarbons (PAH). In utero exposure to particulate matter has been associated with a greater prevalence of PAH-DNA adducts on the placenta, and PAH-DNA adducts, in turn, are correlated with low birth weight, small head circumference, preterm delivery, and fetal deaths (Perera et al. 1998, Topinka et al. 1997, Huel et al. 1993, Hatch et al. 1990). Laboratory experiments on rats have confirmed most of these effects (Ridgon and Rennels 1964, MacKenzie and Angevine 1981). PAHs disrupt central nervous system activity of the fetus, and during critical growth periods such as the third trimester, the disruption has a pronounced effect on fetal growth. PAHs are also hypothesized to reduce nutrient flow to the fetus by suppressing estrogenic and endocrine activity and by binding to placental growth factor receptors (Perera et al. 1999). In utero exposure to PAHs has been linked to increased risk of infant leukemia as well (Alexander et al. 2001).

## 2.2 Description of the Indonesian fires

The 1997 dry season in Indonesia was particularly dry. Figure 1 compares the monthly rainfall recorded at a meteorological station in South Sumatra for 1997 and previous years.

The 1997 dry season was both severe and prolonged: rainfall amounts in June, July, August, and September were lower than usual, and the rainy season was delayed until November. The rest of Indonesia experienced similar rainfall patterns as Sumatra.

Fires are commonly used in Indonesia to clear land for cultivation, and the dry season is considered an opportune time to set fires because the vegetation burns quickly. Industrial farmers burn forest land in order to replant it with palm or timber trees, and small farmers use swiddening or "slash–and–burn" techniques in which land is cleared with fire to ready it for cultivation. In addition, logging companies are thought to have set some virgin forests on fire in order to degrade the land so that the government would designate the land as available for logging.

With expansion of the timber and palm oil industries in Indonesia, many tracts of forestland have become commercially developed, and logged-over land is more prone to fires than pristine forest.[4] Roads running through forests act as conduits for fire to spread, and with the canopy gone, the ground cover becomes drier and more combustible and wind speeds are higher. In addition, because logging firms were taxed on the volume of wood products that left the forest, they often left behind waste wood, even though it had economic value as fertilizer or wood chips. The left-behind debris wood made the forest more susceptible to fast-spreading fires (Barber and Schweithhelm 2000).

In September 1997, because of the dry conditions, the fires spread out of control. The Indonesian government made some attempt to fight the fires, but the efforts were ineffective. The fires continued until the rains arrived in November. In southeastern Kalimantan but not the rest of Indonesia, fires started anew in March 1998 after the rainy season ended.

The fires were concentrated on the island of Sumatra and in Kalimantan. Estimates are that up to 12 million acres burned, 8 million acres in Kalimantan (12% of its land area) and 4 million in Sumatra (4% of its area). The practice of clearing land with fire is used throughout Indonesia, and El Niño affected all of Indonesia. What set Sumatra and Kalimantan apart is that Indonesia's forests are mainly in these areas. The majority of

[4]In 1996 forest products accounted for 10% of Indonesia's gross domestic product, and Indonesia supplied about 30% of the world palm oil market (Ross 2001).

crop plantations are located in Sumatra, and plantations are a fast-growing use of land in Kalimantan. Timber operations are also primarily in these regions.

The location of the smoke generally tracked the location of the fires, though because of wind patterns, not entirely. Figure 2 shows satellite images of the pollution over Indonesia between September and November. Fires were concentrated on the southern parts of Sumatra and Kalimantan, and these two areas experienced the most pollution. On the other hand, the northern half of Sumatra was strongly affected by smoke while Java was relatively unaffected, yet neither of these areas experienced many fires.

A common measure of particulate matter is $PM_{10}$, the concentration of particles less than 10 $\mu$m in diameter. The U.S. Environmental Protection Agency has set a $PM_{10}$ standard of 150 micrograms per cubic meter ($\mu g/m^3$). This is the 24-hour average that should not be exceeded in a location more than once a year. During the 1997 fires, the pollution in the hardest hit areas surpassed 1000 $\mu g/m^3$ on several days and exceeded 150 $\mu g/m^3$ for long periods (Ostermann and Brauer 2001, Heil and Goldmammer 2001).[5] The levels of pollution caused by the wildfires are comparable to levels caused by indoor use of wood-burning stoves. The daily average $PM_{10}$ level from wood-burning stoves, which varies depending on the dwelling and duration of use, ranges from 200 to 5000 $\mu g/m^3$ (Ezzati and Kammen 2002).

# 3 Empirical Strategy and Data

## 3.1 Empirical model and outcome variable

The goal of the empirical analysis is to examine whether air pollution has an effect on fetal or infant death. Ideally, there would be data on all pregnancies indicating which ended in fetal or infant death, and the following equation would be estimated:

$$Survive_{jt} = \beta_1 Smoke_{jt} + \delta_t + \alpha_j + \varepsilon_{jt}. \tag{3.1}$$

[5] One reason the Indonesian fires produced so much pollution is that many were peat fires which produce large amounts of smoke.

The variable $Survive_{jt}$ is the probability that fetuses whose due date is month $t$ and whose mothers reside at the time of the fires in subdistrict $j$ survive to a certain point, such as live birth, one year, etc. The prediction is that $\beta_1$ is negative, or that exposure to smoke around the time of birth reduces the probability of survival.

In practice, mortality records are unavailable for Indonesia, and survey data on infant mortality are not feasible for the analysis because the samples are too small to examine month-to-month fluctuations or geographic variation in pollution. For example, the 2002 Demographic and Health Survey has on average 1 birth and 0.05 recorded child deaths per district-month for the affected cohorts.

Thus, the approach I take is to infer fetal and infant mortality by measuring "missing children."[6] The outcome measure is the cohort size for a subdistrict-month calculated from the complete 2000 Census of Population for Indonesia. The estimating equation is

$$\begin{aligned} ln(CohortSize)_{jt} =& \beta_1 Smoke_{jt} + \beta_2 PrenatalSmoke_{jt} + \\ & \beta_3 PostnatalSmoke_{jt} + \delta_t + \alpha_j + \varepsilon_{jt}. \end{aligned} \tag{3.2}$$

The dependent variable, $ln(CohortSize)_{jt}$, is the natural logarithm of the number of people born in month $t$ who are alive and residing in subdistrict $j$ at the time of the 2000 Census. $Smoke_{jt}$ is the pollution level in the month of birth, and the variables $PrenatalSmoke_{jt}$ and $PostnatalSmoke_{jt}$ are included to explore the different timing of exposure, as discussed below. To obtain parameters that represent the average effect for Indonesia, each observation is weighted by the subdistrict's population (the number of people enumerated in the Census who were born in the year prior to the sample period).

The main advantage of inferring deaths by counting survivors is that the data are for the entire population instead of a sample. Also, the outcome variable measures fetal deaths in addition to infant deaths, albeit without distinguishing between the two outcomes; most surveys do not collect data on fetal deaths. Finally, population counts may be better

[6]The literature on "missing women" in developing countries, most often associated with Sen (1992), uses population sex ratios to infer excess female mortality caused by gender discrimination.

measured than infant mortality because of underreporting of infant deaths and recall error on dates of deaths.

There are several potential concerns about inferring mortality from survivors, however. Since the data come from a cross-section of survivors in June 2000, the outcome represents a different length of survival for individuals born at different times, and the mean level of survival will differ by cohort, independent of the fires. For a cohort born in December 1997 around the time of the fires, the outcome is survival until age two and a half, while for an older cohort born in December 1996, the outcome is survival until age three and a half, for example.[7] The inclusion of birthyear-birthmonth (hereafter, month) fixed effects in the regression will control for any average differences in survival by cohort.

In addition, if pollution affects the duration of pregnancies, then missing children might result from the shifting of births from certain months to other months. For example, if exposure to smoke induces preterm labor, then one would expect to see an excess of births followed by a deficit of births. In section 4.2, I examine and am able to reject the conjecture that the results are an artifact of changes in gestation period.

There are also potential empirical concerns not unique to using $ln(CohortSize)$ as the dependent variable. First, pollution might affect not only mortality but also fertility. This would influence the population counts for the later "control" cohorts and could lead to sample selection problems even if mortality were directly measured. In order that the control cohorts are uncontaminated by fertility effects, I restrict the sample to births occurring no more than eight months after the outbreak of the fires. The last individuals in the sample are those born in May 1998. Second, an implicit assumption in the empirical model is that it is exposure to pollution just before or after birth that affects mortality. The motivation for this model are findings from previous research that exposure to air pollution near the time of birth has significant health effects. However, exposure to pollution earlier in a pregnancy or later after birth also could affect health. If the control cohorts are in fact also

[7] As shorthand I describe deaths of children in the sample as infant mortality even though they could occur as late as age three and a half. The common definition of infant mortality is deaths before age one. Note that one advantage of observing survival more than two years after the due date is that for deaths that occur around birth, the estimates are less likely to reflect simply short-term "harvesting."

treated, though less intensely, then the results would underestimate the true effects.

A third important concern arises from the fact that individuals are identified by their subdistrict of residence in 2000 rather than the subdistrict where their mother resided during the end of her pregnancy or just after giving birth. If families living in high-smoke areas with children born around the time of the fires were more likely to leave the area (either during or after the fires), then cohort size would be smaller in areas more affected by pollution. Fortunately, one can directly examine this concern by analyzing data at the district level since the Census collects the district of birth and the district of residence in 1995. As discussed in section 4.2, the results are identical using birthplace, current location, or mother's location in 1995.

Table 1 presents the descriptive statistics for the sample. The sample comprises monthly observation between December 1996 and May 1998 (18 months) for 3751 subdistricts (*kecamatan*). Of this starting sample size of 67,518 observations, 64 observations are dropped because the cohort size for the subdistrict-month is 0.[8] There are on average 96 surviving children per observation. The larger administrative units in Indonesia are districts (*kabupaten*), of which there are 324 in the sample, and provinces, of which there are 29.

## 3.2 Verification that Census counts track infant mortality

As a preliminary analysis, I verify that population counts from the Census track data on births and infant deaths from the 2002 Demographic and Health Survey (DHS). The log of the number of surviving children should increase one-for-one with the log of total births and should decrease one-for-one with the infant mortality rate (as can be derived with a few steps of algebra). Thus, I estimate

[8]The Census covers 3962 subdistricts which make up 336 districts. For subdistricts dropped from the sample, either the latitude and longitude could not be determined or there were no enumerated children for more than 15% of the monthly observations due to missing data or very small subdistrict size. In addition, I drop four districts that make up Madura since the East Javanese island received a large influx of return migrants in 1999 (in response to ethnic violence against them in Kalimantan), and also Aceh province where separatist violence is thought to have affected the quality of the Census enumeration. The results are also robust to dropping Irian Jaya, another area where unrest could have affected data quality.

$$ln(CohortSize)_{JT} = \alpha + \gamma_1 ln(Births)_{JT} + \gamma_2 IMR_{JT} + \varepsilon_{JT} \tag{3.3}$$

where $J$ is a province and $T$ is a quarter, $Births$ is the number of children born in the province-quarter, and $IMR$ (infant mortality rate) is the fraction of those children who died by June 2000 when the Census was taken. As mentioned above, very few births per subdistrict-month are sampled in the DHS, so the validation exercise aggregates to provinces and quarters and uses a longer panel from 1988 to 1999 to gain power.[9] Note that $ln(Births)$ varies not only with the number of births in the province-quarter but also with the DHS sampling rate for the province. The $IMR$ variable should not be affected by this problem.

Table 2 presents the results of this validation exercise. In column 1, the coefficient on $IMR$ is -1.3 and the coefficient on $ln(Births)$ is 1.6, which are surprisingly close to the predictions of -1 and 1, given the crudeness of the exercise. In column 2, each observation is a province-quarter-gender, and in column 3, a province-month. The coefficients remain on the order of -1 and 1 but become smaller in magnitude, which is consistent with downward bias from measurement error when smaller and hence noisier cell sizes are used. In short, variation in population counts in the Census indeed tracks variation in the number of births and, importantly for this study, variation in the infant mortality rate.

With these results in hand, if one compares equation 3.3 to the estimating equation 3.2, one of the key identifying assumptions becomes apparent. In using $ln(CohortSize)$ as a proxy for the infant mortality rate, in order to obtain unbiased estimates of the effect of pollution on infant mortality, it must be the case that conditional on subdistrict and month fixed effects, pollution is not correlated with $ln(Births)$. This seems like a reasonable assumption. First, by using a short panel, subdistrict fixed effects absorb most variation in the number of women of childbearing age and other determinants of fertility. Month effects control for fertility trends and seasonality. Second, although not observing fertility will

[9]Ideally, the DHS would have recorded pregnancies that ended in fetal deaths. Some of the missing children in the Census are not among the live births measured by the DHS. Also, ideally, the validation exercise would use the same unit of observation and sample period as the main analysis, but the survey data are then too noisy to obtain meaningful results.

add noise to the estimates, it seems unlikely that there were large fluctuations in fertility that coincided with the air pollution both spatially and temporally. Even area-specific trends could not explain the patterns since the sample includes control periods both before and after the fires; any omitted fertility shift causing bias would have to be a short-term downward or upward spike in particular regions. Furthermore, section 4.2 directly tests whether demographic shifts could explain the results, and fluctuations in predicted fertility do not seem to be a confounding factor.

## 3.3 Pollution variable

The measure of air pollution is the aerosol index from the Earth Probe Total Ozone Mapping Spectrometer (TOMS), a satellite-based monitoring instrument. The aerosol index tracks the amount of airborne smoke and dust and is calculated from the optical depth, or the amount of light that microscopic airborne particles absorb or reflect. The TOMS index has been found to quite closely match particulate levels measured by ground-based pollution monitors (Hsu et al. 1999). Ground monitor data are not available for Indonesia for this period. The aerosol index runs from -2 to 7, with positive values representing absorbing aerosols (dust and smoke); for positive values, a higher index indicates more smoke.[10]

The TOMS data set contains daily aerosol measures (which are constructed from observations taken over three days) for points on a 1° latitude by 1.25° longitude grid. Adjacent grid points are approximately 175 kilometers (km) apart. The probe began collecting data in mid-1996, and the data I use begin in September 1996. For each subdistrict, I calculate an interpolated daily pollution measure that combines data from all TOMS grid points within a 100-km radius of the geographic center of the subdistrict, weighted by the inverse distance between the subdistrict and the grid point. The number of TOMS grid points that fall within the catchment area of a subdistrict ranges from 1 and 6 and is on average 4. The mean distance between a subdistrict's center and the nearest grid point is 50 km. The monthly measure is calculated as the median of the daily values, and I also consider

[10]Negative values represent non-absorbing particulates such as sulfates.

the mean of the daily values and the number of days that exceed a (somewhat arbitrary) threshold value of 0.75.

Whereas there are over 3700 subdistricts in the data, there are only 226 unique pollution grid points used. Interpolation adds spatial variation at a finer grain, but uncorrected standard errors would nevertheless overestimate how much independent variation there is in the pollution measure. Moreover, in addition to the non-independence of the pollution variable that arises from interpolation, the actual pollution level is spatially correlated. Therefore I allow for clustering of errors among observations within an island group by month. There are 10 island groups in the sample (Sumatra, Java, Sulawesi, Kalimantan, Bali, West Nusa Tenggara, East Nusa Tenggara, Irian Jaya, Maluku, North Maluku).

The estimating equation (3.2) includes pollution in the month of birth ($Smoke_{jt}$) as well as lags of $Smoke_{jt}$ which measure exposure to pollution in utero, and leads which measure exposure after birth. Note that $Smoke_{jt}$ measures both prenatal and postnatal exposure, with the balance depending on when in the calendar month an individual is born (the Census did not collect the specific date of birth, only the month). It becomes difficult to separately identify each lag and lead with precision, so the main specification uses an average of the pollution level for the three months before the birth month ($PrenatalSmoke_{jt}$) and after the birth month ($PostnatalSmoke_{jt}$). The population-weighted mean values of $Smoke$, $PrenatalSmoke$, and $PostnatalSmoke$ are 0.09, 0.10, and 0.07, as shown in Table 1. On average, the pollution index exceeds 0.75 on 5% of days.

During the months of the fires, September to November 1997, the mean aerosol index for Indonesia was 0.58. For the same months in 1996, the mean was 0.05. Similarly, the mean of $PrenatalSmoke$ was 0.37 for the most affected cohorts (births in October 1997 to February 1998) while during the same months a year earlier, the mean was 0.03. These gaps are helpful when interpreting the magnitudes of the regression coefficients and quantifying the impact of the fires.

The intensity of smoke also varied across Indonesia. Figure 3 shows the average smoke by month for Kalimantan and Sumatra which were the hardest hit regions and for the rest

of Indonesia. Kalimantan, in addition to being the most affected area in 1997, experienced another episode of smoke in early 1998 after the rainy season ended.

## 3.4 Other variables

Several other variables are used in the analysis either as controls or to examine differential effects of pollution, i.e., as interaction terms. First, I construct a measure of the financial crisis that hit Indonesia in late 1997. Cross-sectional variation in the crisis is measured as the 1996 to 1999 ratio of the median log food consumption per capita in a district. The variable is constructed so that it is larger in areas hit harder by the crisis. The consumption data are from the National Socioeconomic Survey (SUSENAS), a large household survey conducted annually by the national statistics bureau. The survey is representative at the district rather than subdistrict level, so data are aggregated to the district. The data appendix describes in more detail how the consumption measure is constructed. The national consumer price index for food is from the central bank and is used as a measure of temporal variation in the crisis. The interaction of these two variables is the crisis measure.

The cross-sectional measure of consumption in 1996 is interacted with the pollution variables to examine how the effects of pollution differ for richer and poorer areas. Measures of the health care system, such as the number of doctors and maternity clinics per capita, as well as the type of fuel people cook with are also used. These variables are from the 1996 Village Potential Statistics (PODES), a census of community characteristics. The PODES has an observation for each of over 66,000 localities which I aggregate to the subdistrict level. In the analyses that use data from the PODES or SUSENAS, the sample size is 63,158 since some Census subdistricts could not be matched to the surveys.

To measure the extent of fires (as opposed to pollution) in an area, daily data on the location of "hot spots" are used. The data are from the the European Space Agency which analyzed satellite measurements of thermal infrared radiation to locate fires. In addition, to control for rainfall I use monthly rainfall totals from the Terrestrial Air Temperature and Precipitation data set and match each subdistrict to the nearest node on the rainfall

data set's 0.5° latitude by 0.5° longitude grid. Finally, I use additional variables from the Census including mother's education and whether a locality is rural or urban.

# 4 Results

## 4.1 Relationship between exposure to smoke and mortality

Table 3, column 1, presents the relationship between cohort size and exposure to smoke. The independent variables are *Smoke*, which is pollution in the month of birth, *PrenatalSmoke* which is pollution in the three months before birth, and *PostnatalSmoke* which is pollution in the three months after birth. The results in column 1 suggest that prenatal exposure to pollution decreases the survival rate of children. *PrenatalSmoke* has a coefficient of -0.035 that is statistically significant at the 1% level. The coefficient for *Smoke* is very close to 0, while the coefficient for *PostnatalSmoke* is -0.014 though statistically insignificant. Standard errors are clustered within an island-month. In column 2, when *PrenatalSmoke* is the only variable in the regression (besides fixed effects), the coefficient is similar to that in column 1.[11] Columns 3 and 4 consider alternative monthly pollution measures, first, the mean rather than median of the daily pollution values and, second, the proportion of days with high pollution (aerosol index above 0.75). Mean pollution gives nearly identical results as the median value, with postnatal exposure now having a negative impact on cohort size that is marginally significant. For the proportion of days with high pollution, the point estimate implies that when there are 3 additional high-smoke days in a month (an increase of 10 percentage points), cohort size decreases by 0.85%.

Exposure to pollution in utero is associated with a decrease in fetal and infant survival. To interpret the magnitude of the effect, note that *PrenatalSmoke* was higher by 0.33 during October 1997 to February 1998 compared to the same calendar months a year earlier; this five-month period are the cohorts for whom *PrenatalSmoke* includes a month

[11] See Table A1 in the appendix for an instrumental variable estimate of the effect of *PrenatalSmoke* on cohort size. The instrument for *PrenatalSmoke* is a dummy for Kalimantan or Sumatra interacted with a dummy for October 1997 to January 1998. The differences-in-differences estimate, which uses only coarse variation in pollution attributable to the fires, is -0.040.

during the fires. Multiplying that gap by the coefficient of -0.035 implies that the fires led to a 1% decrease in cohort size. A more precise way to estimate the total effect is to use the coefficient for *PrenatalSmoke* and calculate what the population would have been for each subdistrict if during the period during and immediately after the fires, *PrenatalSmoke* had taken on its value from 12 months earlier. Aggregated over the five months for the 3751 subdistricts, this calculation similarly implies a population decline of 1.0%, or 16,439 missing children.[12] Indonesia's baseline under-2 mortality rate was roughly 60 per 1000 live births at this time.[13] If the effects of pollution were due exclusively to infant and child deaths, the estimates would represent a 17% effect; if the effects were due in equal part to fetal deaths, the coefficient would imply an 8% effect.

The welfare implications of pollution-induced mortality depend on the counterfactual of how long individuals otherwise would have lived. On average, 1% of children in Indonesia who survive until age 2 die by age 5. Thus, if one wanted to attribute the 1% effect of the fires to "harvesting," essentially all deaths between age 2 and 5 would have to have been pushed forward to the time of the fires. Moreover, by most standards, the shortening of children's lives by even three to five years is a significant welfare loss.

Figure 4 shows the nonparametric relationship between third-trimester exposure and cohort size. The effect of pollution is linear for the most part. There appears to be a somewhat steeper relationship at high levels of pollution, but the data are sparse in this region. The nonlinearities are statistically insignificant when estimated parametrically with a spline or a quadratic term.

The next regressions use the pollution level in each of the three months preceding and following birth, rather than aggregated for a quarter. Table 3, column 5, reports the results using the median pollution level. For prenatal exposure (lags of *Smoke*), the effect is strongest two months before the month of birth. For postnatal exposure (leads of *Smoke*),

[12]The estimates using high-smoke days imply a 1.1% aggregate effect. (The mean of the prenatal high-smoke variable is 0.14 during the 1997-8 episode and 0.01 for the same calendar months a year earlier, and the coefficient in Table 3, column 4, is -0.085.)

[13]The government estimates of under-1 and under-5 mortality rates at this time are 5% and 7%, respectively. I assume that half of deaths between age 1 and age 5 occur before age two.

the effect is strongest immediately after birth, though the estimates are imprecise. The next two columns repeat the exercise using the month's mean pollution and the proportion of days that have high pollution. The general pattern of the point estimates for postnatal pollution remains the same, but the pattern for prenatal exposure is a bit different. For the mean pollution level or number of high-smoke days (columns 6 and 7), exposure in the month immediately before the month of birth now has the strongest negative relationship with cohort size. One interpretation is that at different points during gestation, fetuses are more vulnerable to sustained exposure to pollution versus extreme levels of pollution. A more likely interpretation is that there is not enough precision to determine at this level of detail how the timing of exposure affects survival.[14] Thus, for the rest of the analysis, I focus on the three-month measures of prenatal and postnatal exposure. (The results are similar using two-month measures.)

## 4.2 Effect of smoke on mortality versus alternative hypotheses

The results in Table 3 suggest that exposure to smoke in utero causes infant and fetal deaths. This section considers other possible explanations for the results.

### Migration

The Census identifies respondents by their subdistrict of current residence, but a fetus or infant's exposure to pollution depends on where the family resided during the fires. Migration could be a reason that cohorts with the highest prenatal exposure to pollution are smaller. Women who were in the third trimester of pregnancy during the fires could have been especially likely to migrate away from affected areas, either while pregnant or after giving birth. Fortunately, the Census collects information on the district (though not subdistrict) where an individual was born and where he or she lived five years earlier that enables one to probe this concern.[15]

[14]The month-by-month patterns, unlike the results with the three-month measures, are somewhat sensitive to using a different sample period or a different threshold for high-smoke days, for example.

[15]For 9% of the sample, district of residence differs from district of birth, for 7% of cases, it differs from mother's residence five years earlier, and for 12% of cases, it differs from one or the other.

To examine the extent of pollution-induced migration that occurs *after* birth, I repeat the main analysis by district of birth. Cohort size is aggregated to the district level, and the pollution measure for the district is a population-weighted average of the subdistrict measure. The regression is weighted by the district population in the year preceding the sample period. For comparison, column 1 of Table 4 presents results by district of residence, and column 2 presents results by district of birth. The results are nearly identical to each other, as well as to the subdistrict-level analysis, in terms of both point estimates and precision. Between-district migration after the birth of the infant is not the likely explanation for the relationship between pollution and cohort size.

This finding does not rule out pollution-induced migration that takes place *before* the infant is born. If some women spent most of their third trimester of pregnancy in the hardest-hit areas but migrated away before giving birth, then neither place of residence in 2000 nor place of birth would accurately reflect the fetus' location during the fires. While the Census did not ask respondents where they resided in September to November 1997, it did ask where they lived in 1995. As long as people do not migrate across districts repeatedly, this measure should be a good proxy for where pollution-induced migrants lived at the time of the fires. To test for migration that occurs before birth, I match infants to their mothers as described in the data appendix and repeat the estimate by the district where the mother resided in 1995. The results, shown in column 3, are unchanged from the earlier estimates. In sum, migration, either before or after birth, cannot easily account for the negative relationship between exposure to pollution and cohort size.[16]

**Fertility**

The empirical approach interprets decreases in $ln(CohortSize)$ as increases in early-life deaths, but there would also be fewer survivors if the number of births decreased. It seems unlikely that conceptions declined nine months before the fires with a spatial pattern

[16]Within-district migration is unlikely to be driving the results since there is very little within-district variation in pollution, and most of it derives from interpolation so is noisy. In a model with district-month fixed effects, the coefficient for $PrenatalSmoke$ is -0.013, smaller than in the main specification (Table 3, column 1), and is imprecise, suggesting that between-district variation is dominant in the main estimates.

matching the pollution, but this omitted variable concern also can be tested more directly. To do so, I construct a measure of predicted births. First, I measure the percentage of women of each age who give birth, using a time period not in the sample (namely, the youngest cohorts in the Census, those born in 1999 and 2000, so that survivors most closely approximate births). I then apply these birth rates to the demographic composition of each district-month in the sample. This gives a predicted number of births based on demographic shifts. (See the data appendix for further details.) Table 5, column 1, shows the results when $ln(PredictedBirths)$ is included as a control variable. The coefficient of survivors on births is predicted to be slightly less than 1. Because the measure is noisy especially after conditioning on subdistrict and month indicators, the estimate is likely to suffer from attenuation bias. The estimated coefficient on predicted births is less than but statistically indistinguishable from 1. More importantly, the coefficients on the pollution variables are essentially unchanged with this control variable included. Fluctuations in fertility, at least those caused by demographic shifts, do not appear to be a confounding factor in the analysis.[17]

**Preterm births**

Another hypothesis is that the missing children are not deaths but instead reflect changes in the duration of pregnancies. In particular, exposure to pollution may have induced preterm delivery which is often associated with traumatic pregnancies. The reason this mechanism could conceivably generate the results is that it is *prenatal* exposure that has a strong negative relationship with cohort size. Consider September 1997, the month the fires started. Pollution levels were high in September but the value of $PrenatalSmoke$ in September is low since there was no significant smoke in June, July, or August. In October, in contrast, $PrenatalSmoke$ is high since it incorporates the pollution in September. If infants due in October were instead born in September, then births would have shifted

[17]Table A2 addresses another potential concern about fertility, namely that the seasonality of births or deaths could happen to differ for areas more affected by the pollution, generating a spurious result. As shown in columns 4 and 5, the results are robust to restricting the sample to the months with high $PrenatalSmoke$ plus the same calendar months one year earlier.

from a high-*PrenatalSmoke* month to a low-*PrenatalSmoke* month, generating a negative relationship between *PrenatalSmoke* and cohort size that is unrelated to mortality.[18] To test the preterm-birth hypothesis, I repeat the analysis excluding September 1997 from the sample. If the above hypothesis were correct, the coefficient on *PrenatalSmoke* would become less negative (and the coefficient on *Smoke* would become more negative) compared to the baseline results. As shown in Table 5, column 2, this does not occur. The estimated coefficients are nearly identical between the full sample and the subsample, contrary to what one would expect if the pollution had induced preterm births but had not affected infant survival.[19]

According to the results of this test, the effect of pollution on cohort size is not due to preterm births *instead of* being due to fetal and infant deaths. However, pollution may have caused infant deaths precisely by inducing premature births (which put infants at greater risk of death); that is, preterm delivery is potentially an important channel through which exposure to pollution led to mortality.

**Financial crisis**

The Indonesian financial crisis began shortly after the 1997 fires, as shown in Figure 5, so a concern is that the analysis is attributing to air pollution deaths that were caused by the crisis. To test this alternative hypothesis, a measure of the financial crisis is added to the model. No monthly subdistrict-specific data on the crisis were collected, to my knowledge, so I construct a measure of the crisis by interacting a cross-sectional measure, the inverse ratio of median income (consumption) at the height of the crisis in 1999 to median income before the crisis in 1996, and a time-series measure, the consumer price index for food. The regression results can be anticipated by noting that the cross-sectional correlation between the crisis measure and pollution in October 1997 (peak of the fires) is

[18]Considering only September, this phenomenon should also generate a positive correlation between *Smoke* (pollution in the month of birth) and cohort size, but averaged with October and November which have high values of both *Smoke* and *PrenatalSmoke*, the net effect is indeterminate.

[19]Table A3 restricts or expands the sample to other time periods, and the results are robust to this change. One noteworthy finding is that the estimated effect of *PrenatalSmoke* is smaller when the window extends more than 8 months after the fires, suggesting that the fires may have reduced fertility.

0.04; the spatial patterns of the crisis are not similar to the spatial patterns of pollution from the wildfires. For regressions that use variables from the SUSENAS or PODES surveys, a slightly smaller sample of subdistricts is used due to data availability. Table 5, column 3, shows the regression results for the baseline model without additional regressors and confirms that the subsample is similar to the full sample. Columns 4 shows the results when the crisis variable contemporaneous to the month of birth is included as a control variable. The estimated effect of *PrenatalSmoke* remains -0.032. The crisis measure has been normalized to have a mean of zero and standard deviation of one for the sample, so the coefficient implies that a one standard deviation increase in the crisis is associated with a 4.9% smaller cohort, though the coefficient is statistically insignificant (and moreover could be due to migration rather than mortality).[20,21]

**Effect of pollution versus effect of fires or drought**

Another interpretation of the results is that they represent reduced-form mortality effects of the fires rather than effects of specifically air pollution from the fires. The regressor is the pollution level, and previous research gives one reason to expect that pollution causes infant mortality, but, by and large, the smoke affected places nearby the sites of fires, and the fires could have caused mortality through income effects, degraded food supply, and other channels. To separate the effect of pollution from other effects of the fires, I use data on where precisely the fires occurred. I calculate the number of fire-days occurring in or near a subdistrict based on satellite data on "hot spot" locations and durations. Fire-days is the duration of each fire summed over all fires within 50 km of the subdistrict center. First, I examine the effects of pollution in areas that did not experience extensive fires. In Table 5, column 5, the sample is restricted to subdistricts where fewer than 80 fire-days occurred over the sample period (which eliminates 10% of subdistricts, predominantly in Kalimantan and Sumatra). The coefficient on *PrenatalSmoke* on log cohort size remains -0.035 for these

[20]Since the crisis accelerated a few months after the fires, I also estimated models that control for the crisis measure for the three months following the month of birth. This generates more variation in the crisis measure during the period of interest. The estimated effect of *PrenatalSmoke* remains unchanged.

[21]Rukumnuaykit (2003), using the Indonesia Family Life Survey, finds a 3% increase in infant mortality in 1997-8 which is interpreted as due to the financial crisis as well as to the drought and smoke.

areas which experienced the pollution from the fires but not the fires themselves. Next, I include measures of fire prevalence as regressors. The number of fire-days is a highly skewed variable, so I use two indicator variables, one for whether there were any fire-days in the subdistrict-month (sample mean of 0.16) and a second for whether there were intense fires, defined as at least 10 fire-days during the month (sample mean of 0.03). In column 6, the fires variable and intense fires variable in the month of birth, averaged over the three months before birth (prenatal exposure), and averaged over the three months after birth (postnatal exposure) are included as regressors. The effect of $PrenatalSmoke$ is -0.032, nearly identical to earlier estimates, which supports the interpretation that air pollution is the cause of the increase in fetal and infant mortality. There is also some evidence that intense fires in the month of birth are associated with a decrease in cohort size, suggesting that fires may have an additional effect on survival (or migration) through channels besides pollution. The effect size is relatively small. The intense fires variable is larger by 0.09 during the 1997 episode compared to a year earlier, so the coefficient of -0.028 implies that intense fires are associated with a 0.25% decrease in cohort size.

Another hypothesis is that the effects are due not to pollution but to drought. There was below-normal rainfall throughout Indonesia in 1997, not just in areas affected by pollution, so, conditional on month effects, drought seems unlikely to account for the findings. The hypothesis also can be tested directly by controlling for rainfall. Specifically, I control for rainfall four months prior to birth. This lag was chosen because it stacks the cards in favor of finding that rainfall is driving the effects; the especially dry months of June to September 1997 are allowed to affect the cohorts found to be most affected by pollution, namely those in the late third trimester during the September to November fires, or those born from October to January. Monthly rainfall for the subdistrict is measured relative to the 1990–5 average for that calendar month. The variable's sample mean is 0.86, and the mean for the October 1997 to January 1998 cohorts is 0.43. As reported in Table 5, column 7, the coefficients for the pollution variables are largely unaffected when rainfall is included as a

control. The coefficient for rainfall is statistically insignificant and quantitatively small.[22] The changes in cohort size do not seem to be due to rainfall shortages.

## 4.3 Effects by gender and income

### Gender

This section examines how the mortality effects of pollution vary across groups. I first test whether there are differential effects for boys and girls. Table 6, column 1, reports results for a model in which the number of surviving boys and girls are totaled separately, each observation is a subdistrict-month-gender, and the three pollution variables are interacted with a dummy for male. The coefficients are imprecise but follow an interesting pattern. The male interaction terms are positive for contemporaneous and postnatal smoke, but negative for prenatal smoke. The more negative effect for boys in utero (30% larger effect) is consistent with findings in the literature that male fetuses are less physiologically robust than female fetuses and in particular have more delayed lung development (Hassold, Quillen, and Yamane 1983, Jakobovits 1991). The more negative effect for girls in the month of birth and after birth could reflect physiological differences, but it is also consistent with gender discrimination. If parents are more likely to take steps to minimize a newborn son's exposure to pollution or to seek medical treatment for his respiratory infection, for example, then one would expect the effects of postnatal pollution to be stronger for girls.[23]

### Income

The next estimates test whether the effects of pollution are more pronounced in poorer places. People in poorer areas might have lower baseline health, making them more sensitive

[22]The results are similar using contemporaneous rainfall or rainfall with different lags. In particular, one concern is that drought might reduce fertility. This is unlikely to explain the effects since the drought began only in June, but if rainfall shortages began in May in some areas, then cohorts conceived in May were entering the third trimester in November and could be affected in this way. When rainfall nine months prior to birth is added as a control variable, again, rainfall has a small and insignificant effect on cohort size and the effect of prenatal pollution is unchanged.

[23]If the health effects of prenatal exposure assert their symptoms after birth and parents favor boys, one would expect boys to also be less affected by prenatal exposure. This effect on survival would offset any physiological disadvantage that male fetuses have when exposed to pollution prenatally.

to pollution; have less access to health care; or have effectively more exposure to pollution because they spend more time outdoors or performing strenuous tasks and breathing heavily, for example. Column 2 of Table 6 uses food consumption as a proxy for income to examine this hypothesis, interacting the pollution measures with a dummy variable for whether the subdistrict's median log consumption in 1996 is above the 50th percentile among all subdistricts. All three of *Smoke*, *PrenatalSmoke*, and *PostnatalSmoke* are associated with smaller cohorts for the bottom half of the consumption distribution, and the interaction terms for the top half of the distribution are large and positive. The model estimated in column 2 appears to be misspecified, however. The weighted average of the coefficients for the bottom and top halves of the distribution would be more negative than the average effect found earlier. The reason for the apparent paradox is that month effects vary significantly with income. As has been documented in the demography literature, seasonality in fertility tends to be stronger and qualitatively different in poorer areas (Lam and Miron 1991). Thus, column 3 includes separate month fixed effects for the top and bottom halves of the consumption distribution. The results are qualitatively similar to those in column 2. The effect of prenatal exposure is large and negative when consumption is below the median. In these areas, postnatal exposure is also statistically significant, with an effect size about 60% that of prenatal exposure. Each of the interaction coefficients for districts with above median consumption is positive, and in the case of *PrenatalSmoke*, significant at the 1% level. The effect of a one unit change in *PrenatalSmoke* is -0.06 for the top half of the distribution and -0.13, or over twice as large, for the bottom half. Average log consumption is 0.4 log points larger in the top half of the distribution compared to the bottom half, so another way to view the results is that when consumption increases by 50% ($e^{0.4}$), the effect size decreases by 50%.

The fact that seasonal patterns in cohort size differ by income suggests that including separate month effects for the two halves of the consumption distribution might be the preferred specification even for estimating the average effect. As shown in column 4, the average effect for prenatal smoke is then -0.069 and the coefficient for postnatal smoke is

-0.033, both twice as large as seen earlier in Table 3.

Finally, I further break down the income distribution into quartiles (and include month-quartile fixed effects). Column 5 shows the separate coefficients by quartile, estimated as one regression. The point estimate on *PrenatalSmoke* becomes more negative moving from higher to lower quartiles. The results are not very precise, though, and the *PrenatalSmoke* coefficients for different quartiles are not statistically distinguishable from one another. The coefficients for the other smoke variables are also imprecise, especially for the bottom two quartiles, and the point estimates do not monotonically decline with consumption. Above- versus below-median consumption, as opposed to a linear interaction term, is therefore used below to parsimoniously characterize the heterogeneous effects by income.

**Effects by urbanization, wood-stove use, health care, and mother's education**

There are several possible reasons for the income gradient in the effects of pollution, and this subsection tests some hypotheses. The evidence presented here is merely suggestive since the measures used could be correlated with omitted variables and since data are available to test only a limited number of hypotheses.

One possibility is that urban areas experience smaller effects from the fires than rural areas, and it is this fact that generates the heterogeneity by income. Urbanization would only be a proximate cause, but one might think that in urban areas, housing stock is less permeable to pollution, health care is better, there is less outdoor work, or there are more effective public advisories urging people to stay indoors, for example. On the other hand, pollution from the fires may have been particularly noxious in cities where it mixed with industrial pollution from cars and factories. Column 1 of Table 7 interacts the pollution measures with the proportion of the subdistrict population that lives in urban localities (based on those born in the year before the sample period). Only the coefficients for *PrenatalSmoke* and its interaction terms are reported, but *Smoke*, *PostnatalSmoke* and their interactions are also included in the regressions. The effects of pollution do not vary by urbanization level, suggesting that the offsetting effects described above may have cancelled

each other out.[24]

Next I test whether the effects depend on the prevalence of wood-burning stoves. If the health impact of pollution is convex in exposure, those who have daily exposure to indoor air pollution could suffer more acutely from the wildfires. For each village or town, data are available on whether the majority of people used wood or other biomass as their primary cooking fuel in 1996. I construct for each subdistrict the population-weighted average of this measure, which serves as a crude measure of the proportion of people in the subdistrict who use wood as their cooking fuel. The mean of the variable is 0.64. As shown in column 2 of Table 7, wood fuel use is strongly associated with more negative effects from any given level of exposure to the wildfire smoke. A 20 percentage point decrease in wood fuel use reduces the net effect of prenatal pollution by 0.03. By comparison, moving from the bottom half to the top half of the consumption distribution (50% increase in consumption) reduces the effect of prenatal pollution by 0.07. The obvious caveat to these results is that use of wood-burning stoves could be proxying for an omitted variable.

I also examine whether the effects vary with the availability of health care in the area. A good health care system could lead to improved baseline health of mother and child through prenatal care and to better medical treatment of morbidities caused by the pollution, for example. Table 7, columns 3 to 4, present the results when interactions of the pollution measures with maternity clinics and doctors per capita are successively included. The per capita measures, which are for 1996, have been normalized to be mean 0, standard deviation 1. In areas with more maternity clinics or doctors, pollution has a significantly smaller effect on cohort size. The net coefficient for $PrenatalSmoke$ is smaller by 0.03 in an area with one standard deviation above the average number of maternity clinics compared to the average area, and smaller by 0.05 in an area with one standard deviation above average doctors per capita. These results are similar to the those of Frankenberg (1995). She examines within-village changes in the health sector between 1983 and 1986 and finds that infant

[24]In unreported results, when the sample is divided into infants born to mothers who work in agriculture, work in other industries, or do not work, it does not appear that children whose mothers work in agriculture experience larger effects.

mortality decreases when a village acquires more maternity clinics or doctors.

In column 5, the interactions with wood fuel use, doctors, and maternity clinics are estimated in a single regression. The effects of in utero pollution continue to be considerably larger when wood fuel use is higher. In addition, the interaction terms for maternity clinics and doctors remain positive and, for maternity clinics, statistically significant, though smaller in magnitude than when estimated separately. What is as noteworthy is that these interaction terms do not fully explain the differential effects by income.

Table 8 examines how the effects vary by mother's education. Mother's education is important per se as a factor that might affect children's health, and it is also the best available individual-level (i.e., Census) measure of socioeconomic status. I match each infant to his or her mother and create an indicator variable for whether the mother has completed junior high. The sample mean is 0.38. In column 1, there are two observations for each subdistrict-month, the number of surviving children for high-education and for low-education mothers. The three pollution variables are interacted with an educated-mother dummy. The main effects for both prenatal and postnatal pollution are negative, sizeable, and statistically significant. The interaction terms are noisily estimated, but the point estimates are quite striking: they are positive and the same magnitude as the main effects, suggesting that the mortality effects of pollution are confined to children of less educated mothers. Next I examine whether differences in maternal education across areas explain why the effects of pollution are smaller in places with higher consumption. In column 2, the pollution variables are interacted with, first, the percentage of children in the subdistrict born to educated mothers and, second, a dummy for above-median consumption. The interaction of $PrenatalSmoke$ with income remains positive and statistically significant, while the interaction with mother's education does not. Income rather than maternal education seems to be the component of socioeconomic status that dampens the effects of pollution in more developed areas. Note that for $PostnatalSmoke$, both interaction coefficients remain positive, and it is the interaction with maternal education that is marginally significant.

The factors examined in this section are unable to fully explain why prenatal exposure

to pollution has a smaller effect on mortality in higher income areas. Even after controlling for wood-stove pollution, access to health care, and maternal education, there remains a positive coefficient for the interaction of pollution and high consumption. With better measures of indoor pollution, health care, and parental education, one might be able to explain more of the heterogeneity by income. There are also several channels not tested here. Mothers and newborns in poor areas might be less healthy to begin with because of poor nutrition. Dwellings in poor areas might be more permeable to smoke, allowing pollution from the wildfires to contaminate indoor air quality. Behavioral responses might differ by income, with those in richer areas being more likely to stay indoors, avoid strenuous activity or temporarily evacuate to less affected areas, for example.[25] But regardless of the underlying reason, the differential effects by income suggest that the mortality costs of pollution are disproportionately borne by the poor.

# 5 Conclusion

Air pollution from the land fires that engulfed Indonesia in late 1997 caused over 16,400 infant and fetal deaths, or a 1 percentage point decrease in survival for the affected cohorts. This paper exploits the abrupt timing of the pollution and the spatial variation across Indonesia to identify these effects. The paper presents evidence on which timing of exposure matters most: in utero exposure to pollution has the largest effect on survival. At levels that are common both indoors and outdoors in many poor countries, particulate matter has a sizeable effect on early-life mortality.

Questions in developing countries sometimes go unstudied because data are not available. Research on infant mortality effects of air pollution in the United States makes use of

[25] In unreported results, the effects of prenatal smoke are smaller in areas where more houses are constructed with bricks and concrete (compared to wood, palm leaves, etc.). Worse road quality, which could be a proxy for higher costs of evacuating, is also associated with larger mortality effects of pollution. Another approach to measuring evacuation would be to use distance to the nearest low-smoke area, but in practice this variable is too highly correlated with the pollution level to be useful as an interaction term. Use of surgical masks might also vary with income. Kunii et al. (2002) surveyed 532 people during the fires and found that use of surgical masks was associated with fewer respiratory problems, but most medical experts believe that surgical masks are ineffective at blocking out fine particles.

linked natality-mortality records and ground-based pollution monitors. No such data exist for Indonesia. To overcome this obstacle, this paper uses an unconventional methodology. First, infant and fetal death are inferred from "missing children" in the 2000 Indonesian Census. While the indirect method could introduce potential problems, the paper is able to show that migration, changes in gestation length, and other potential concerns do not seem to be driving the results. Second, smoke data from a satellite-based spectrometer are used in lieu of ground-monitor pollution data. Because of the satellite's global coverage, proxy measures of particulate matter are available for even remote, underdeveloped areas.

There are at least two broader lessons about environmental issues in developing countries worth highlighting. First, environmental damage—and the accompanying health effects—are yet an additional consequence of weak governance. Corruption, which is prevalent in Indonesia as in many low–income countries, was an important factor behind the catastrophic fires. The Suharto government turned a blind eye when large firms started fires in violation of the law. One man trying to hold firms accountable was the Minister of Forestry. In September 1997 he named 176 firms suspected of illegally setting fires. However, in a move that was brazen even by Suharto's formidable standards of crony capitalism, in early 1998 he appointed his golfing partner Bob Hasan—a timber magnate—as the new Minister of Forestry. Hasan was outspoken in blaming small farmers for the fires and exonerating large firms, including his own. In virtually no cases were firms punished for starting illegal blazes.

The findings of the paper highlight a second link between the environment and economic development: the health burden from pollution seems to fall disproportionately on the poor. The estimated effect size for fetal and infants deaths is much larger in poorer areas than in richer areas. There is suggestive evidence that the heterogeneity could be because people in underdeveloped areas use wood-burning stoves and face a compounded effect of indoor plus outdoor air pollution. Part of the explanation also may be less access to health care and lower parental education. For the most part, though, why the health effects of pollution vary with income is an open question—and an area to pursue to better understand how environmental degradation creates unique challenges in developing countries.

# Data Appendix

**Census data**
Indonesia conducted a Census of its population in June 2000. The dependent variable, the cohort size for a subdistrict-month, is the count of all enumerated people born in a given month who reside in the subdistrict. The specific date of birth is not available. The population weight for each subdistrict is the total number of people born in 1994 to 1996 enumerated in the Census.

I link mothers to children for the analyses by mother's 1995 residence and education and for constructing the measure of predicted births. Using a household identifier, I link each child to women who are 14 to 42 years older than the child. When there are multiple matches, I give preference to household heads or spouses of heads and to women closer to the peak of the fertility age distribution. To construct predicted fertility, I perform this matching for children born in 1999 and 2000 (through May) and calculate the mother-child age gap. Then for each age in months of women between 14 and 42, I calculate the number of children these women give birth to divided by the total number of women of that age. This gives the fertility rate (net of infant mortality) for each age. I make two adjustments to the fertility rate. First, I smooth the distribution using values for the 4 ages in months before and after each data point. Second, after the age of 38, I replace the estimate with a linear extrapolation from the estimated value at age 38 to 0 at age 42 to correct for the fact that the matching process mistakenly assigns grandmothers as mothers in some cases (giving an implausibly high fertility rate for older ages). The next step is to calculate the number of women by age for each district in the sample period, and multiply it by the age-specific fertility rate. Summing across all the ages of women of childbearing age gives the predicted number of births for each district-month in the sample.

**TOMS pollution data**
In addition to the information provided in the text, further details on the Total Ozone Mapping Spectrometer can be found at http://toms.gsfc.nasa.gov.

**Fire location data and rainfall data**
The European Space Agency compiles the ATSR World Fire Atlas, a list of "hot spots" (date, time, latitude and longitude) identified by nighttime infrared measurements by the ATSR-2 instrument onboard the ERS-2 satellite. A hot spot corresponds to a reading of $> 312$ Kelvins at 3.7 micron wavelength. For each subdistrict-month, hot spot-days within 50 km of the subdistrict's center are used to calculate the number of fire-days, or the sum over discrete fires of the duration of the fire in days.

The rainfall data set, Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series, is from the Center for Climatic Research, University of Delaware. The rainfall measure for each node on a $0.5^\circ$ latitude by $0.5^\circ$ longitude grid is interpolated from 20 nearby weather stations using a spherical version of Shepard's distance-weighting method. The rainfall measure for a subdistrict uses the closest node.

**SUSENAS data**
I use household level data from the 1996 and 1999 SUSENAS core modules which aggregate item-by-item consumption data to two categories, food and non-food. For each household, per capita consumption is calculated weighting children by 0.75 and infants by 0.6 compared to adults who are weighted by 1. The subdistrict log consumption measure is the median across all households of log food consumption per capita.

**PODES data**
The PODES is a census of all villages and towns in Indonesia. I use the population, fuel use, and health facilities questions for 1996. One question asks what cooking fuel the majority of the village uses, where I group the answers as wood fuel (wood plus other biomass) or other (kerosene and gas). The population-weighted average of this indicator variable across villages in a subdistrict is the fuel use variable. Health care measures are unweighted per capita measures for the subdistrict, based on the reasoning that people have access to facilities throughout the subdistrict.

# References


Alexander, F., S. Patheal, A. Biondi, and Others (2001): "Transplacental Chemical Exposure and Risk of Infant Leukemia with MLL Gene Fusion," *Cancer Research*, 61(6), 2542–6.

Barber, C., and J. Schweithhelm (2000): *Trial by Fire: Forest Fires and Forestry Policy in Indonesia's Era of Crisis and Reform*. World Resources Institute, Washington, DC.

Berkowitz, G., M. Wolff, T. Janevic, I. Holzman, R. Yehuda, and P. Landrigan (2003): "The World Trade Center Disaster and Intrauterine Growth Restriction," *Journal of the American Medical Association*, 290(5), 595–6.

Bobak, M., and D. Leon (1992): "Air Pollution and Infant Mortality in the Czech Republic, 1986–1988," *Lancet*, 340, 1010–1014.

Chay, K., and M. Greenstone (2003a): "Air Quality, Infant Mortality, and the Clean Air Act of 1970," National Bureau of Economic Research Working Paper No. 10053.

——— (2003b): "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), 1121–1167.

Currie, J., and M. Neidell (2005): "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?," *Quarterly Journal of Economics*, forthcoming.

Dejmek, J., S. Selevan, B. Benes, I. Solansky, and R. Sram (1999): "Fetal Growth and Maternal Exposure to Particulate Matter During Pregnancy," *Environmental Health Perspectives*, 107, 475–480.

Emmanuel, S. (2000): "Impact to Lung Health of Haze from Forest Fires: The Singapore Experience," *Respirology*, 5(2), 175–182.

Ezzati, M., and D. Kammen (2002): "The Health Impacts of Exposure to Indoor Air Pollution from Solid Fuels in Developing Countries: Knowledge, Gaps, and Data Needs," *Environmental Health Perspectives*, 110(11), 1–12.

Frankenberg, E. (1995): "The Effects of Access to Health Care on Infant Mortality in Indonesia," *Health Transition Review*, 5, 143–163.

Frankenberg, E., D. McKee, and D. Thomas (2004): "Health Consequences of Forest Fires in Indonesia," *Demography*, 42(1), 109–129.

Friedman, M., K. Powell, L. Hutwagner, L. Graham, and W. G. Teague (2001): "Impact of Changes in Transportation and Commuting Behaviors During the 1996 Summer Olympic Games in Atlanta on Air Quality and Childhood Asthma," *Journal of the American Medical Association*, 285, 897–905.

Hassold, T., S. Quillen, and J. Yamane (1983): "Sex Ratio in Spontaneous Abortions," *Annals of Human Genetics*, 47, 39–47.

Hatch, M., D. Warburton, and R. Santella (1990): "Polycyclic Aromatic Hydrocarbon-DNA Adducts in Spontaneously Aborted Fetal Tissue," *Carcinogenesis*, 11, 1673–5.

Heil, A., and J. Goldmammer (2001): "Smoke-Haze Pollution: A Review of the 1997 Episode in Southeast Asia," *Regional Environmental Change*, 2(1), 24–37.

Her Majesty's Public Health Service (1954): "Mortality and Morbidity During the London Fog of December 1952," Reports on Public Health and Medical Subjects.

Hsu, N., J. Herman, O. Torres, B. Holben, D. Tanre, T. Eck, A. Smirnov, B. Chatenet, and F. Lavenu (1999): "Comparisons of the TOMS Aerosol Index with Sunphotometer Aerosol Optical Thickness: Results and Applications," *Journal of Geophysical Research*, 104, 6269–6280.

Hueglin, C., C. Gaegauf, S. Kunzel, and H. Burscher (1997): "Characterization of Wood Combustion Particles: Morphology, Mobility and Photoelectric Activity," *Environmental Science and Technology*, 31, 3439–3447.

Huel, G., J. Godin, N. Frery, F. Girard, T. Moreau, C. Nessmann, and P. Blot (1993): "Aryl Hydrocarbon Hydroxylase Activity in Human Placenta and Threatened Preterm Delivery," *Journal of Exposure Analysis and Environmental Epidemiology*, 3 (suppl. 1), 187–9.

Jakobovits, A. (1991): "Sex Ratio of Spontaneously Aborted Fetuses and Delivered Neonates in Second Trimester," *European Journal of Obstetrics and Gynecology and Reproductive Biology*, 40, 211–213.

Kunii, O., S. Kanagawa, I. Yajima, Y. Hisamatsu, S. Yamamura, T. Amagai, and I. Ismail (2002): "The 1997 Haze Disaster in Indonesia: Its Air Quality and Health Effects," *Archives of Environmental Health*, 57, 16–22.

Lam, D., and J. Miron (1991): "Seasonality of Births in Human Populations," *Social Biology*, 38(1-2), 51–77.

Loomis, D., M. Castillejos, D. Gold, W. McDonnell, and V. Borja-Aburto (1999): "Air Pollution and Infant Mortality in Mexico City," *Epidemiology*, 10, 118–123.

MacKenzie, K., and D. Angevine (1981): "Infertility in Mice Exposed In Utero to Benzo[a]pyrene," .

Ostermann, K., and M. Brauer (2001): "Air Quality During Haze Episodes and Its Impact on Health," in *Forest Fires and Regional Haze in Southeast Asia*, ed. by P. Eaton, and M. Radojevic, chap. 10. Nova Science, Huntington, NY.

Perera, F., R. Whyatt, W. Jedrychowski, V. Raugh, D. Manchester, R. Santella, and R. Ottman (1998): "Recent Developments in Molecular Epidemiology: A Study of the Effects of Environmental Polycyclic Aromatic Hydrocarbons on Birth Outcomes in Poland," *American Journal of Epidemiology*, 147, 309–314.

Perera, F. P., W. Jedrychowski, V. Rauh, and R. M. Whyatt (1999): "Molecular Epidemiologic Research on the Effects of Environmental Pollutants on the Fetus," *Environmental Health Perspectives Supplements*, 107(suppl. 3), 451–60.

Pope, C. A., J. Schwartz, and M. Ransom (1992): "Daily Mortality and PM10 Pollution in Utah Valley," *Archives of Environmental Health*, 15, 3–23.

Rabindran, S. (2001): "Essays in Empirical Environmental Economics: GIS–Econometric Analysis of Indonesia's Fires, Bolivia's Deforestation and Mexico's Trade with the United States," Ph.D. thesis, Massachusetts Institute of Technology.

Ridgon, R., and E. Rennels (1964): "Effect of Feeding Benzpyrene on Reproduction in the Rat," *Experientia*, 4, 224–6.

Robinson, P. (1988): "Root-n Consistent Semiparametric Regressions," *Econometrica*, 56, 931–954.

Ross, M. L. (2001): *Timber Booms and Institutional Breakdown in Southeast Asia*. Cambridge University, Cambridge, UK.

Rukumnuaykit, P. (2003): "Crises and Child Health Outcomes: The Impacts of Economic and Drought and Smoke Crises on Infant Mortality and Birthweight in Indonesia," Mimeo, Michigan State.

Sastry, N. (2002): "Forest Fires, Air Pollution, and Mortality in Southeast Asia," *Demography*, 39(1), 1–23.

Sen, A. (1992): "Missing Women," *British Medical Journal*, 304, 586–87.

Tacconi, L. (2003): "Fires in Indonesia : Causes, Costs and Policy Implications," Discussion Paper No. 38, Center for International Forestry Research (CIFOR).

Topinka, J., B. Binkova, G. Mrackova, and Others (1997): "Influence of GSTM1 and NAT2 Genotypes on Placental DNA Adducts in an Environmentally Exposed Population.," *Environmental and Molecular Mutagenics*, 30(2), 184–95.

Wang, X., H. Ding, L. Ryan, and X. Xu (1997): "Association between Air Pollution and Low Birth Weight: A Community-Based Study," *Environmental Health Perspectives*, 105, 514–520.

Figure 1: Rainfall at Palembang Airport meteorological station, South Sumatra, 1990-97

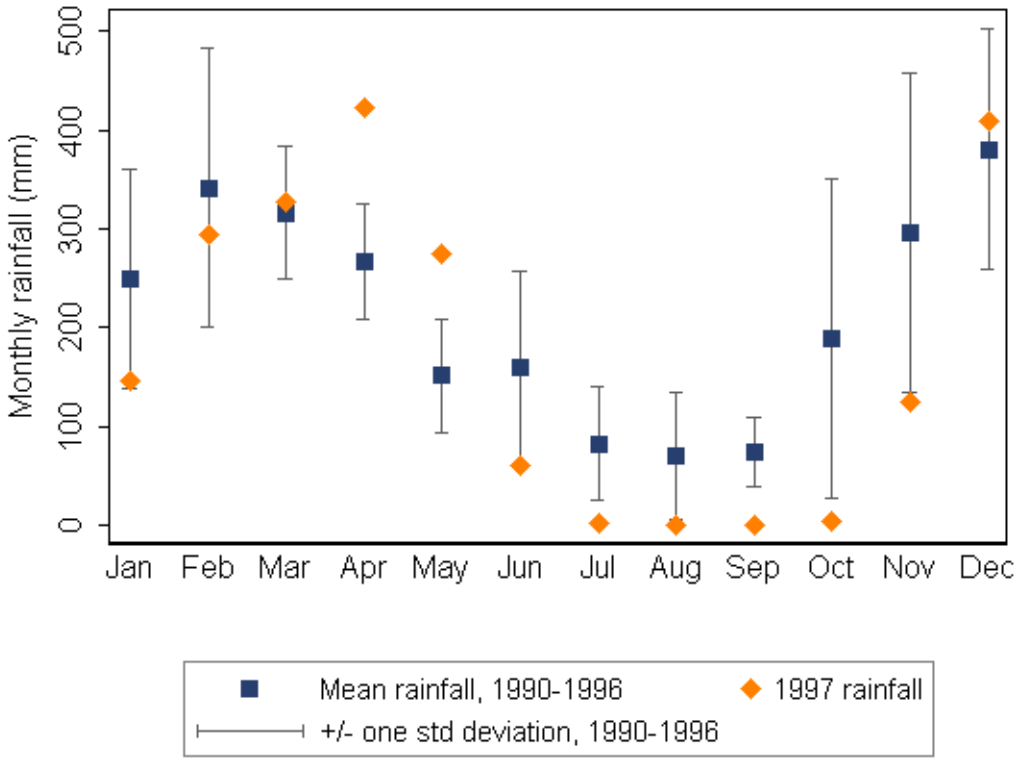
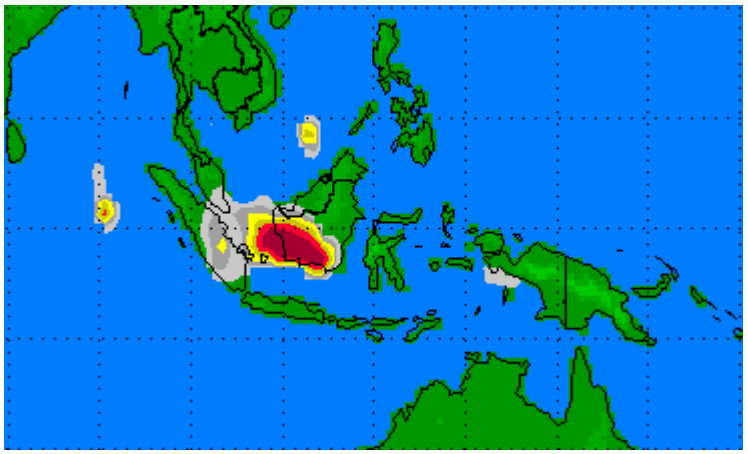
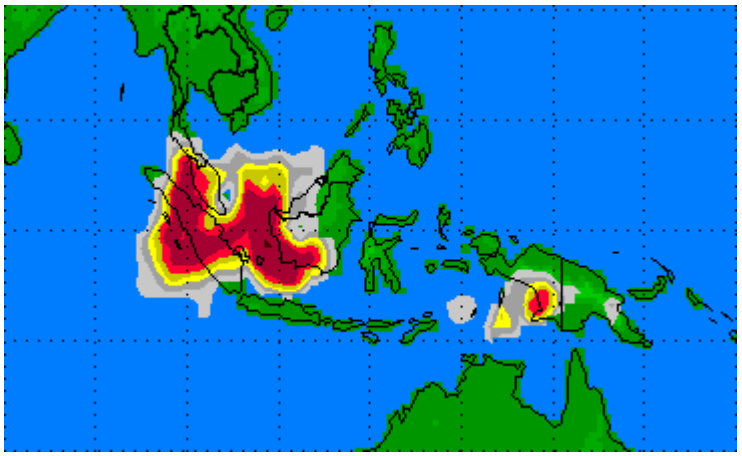


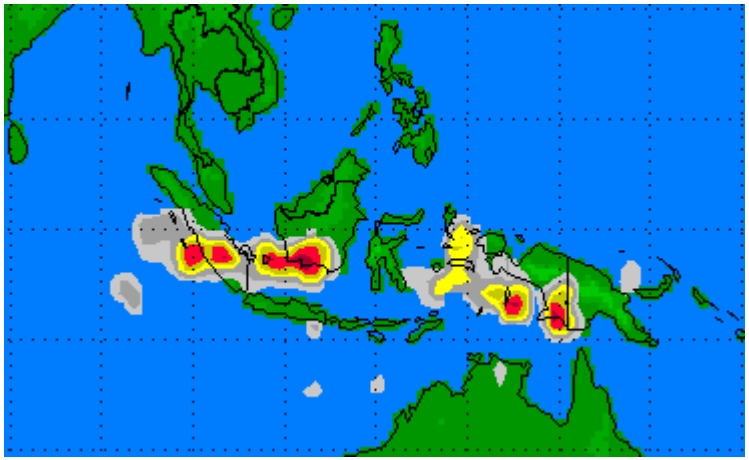
Figure 2: Satellite images of smoke over Indonesia



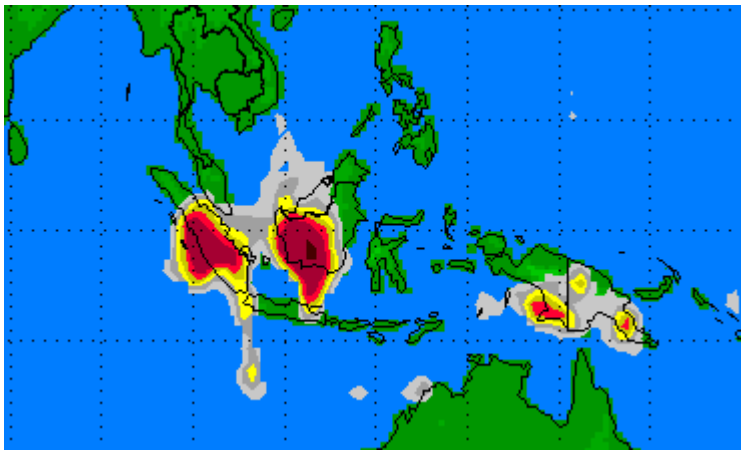
September 5, 1997



September 25, 1997



October 15, 1997



November 4, 1997

Figure 3: Timing and location of the pollution

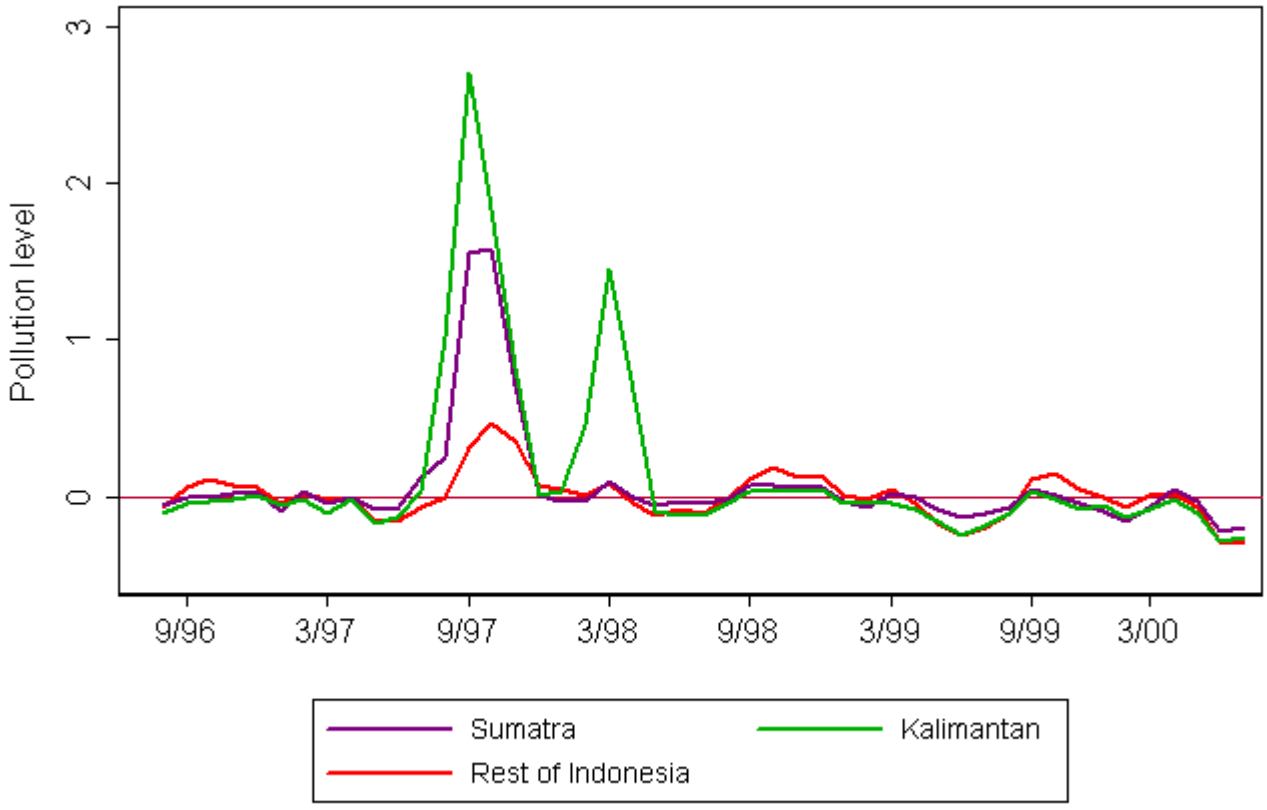
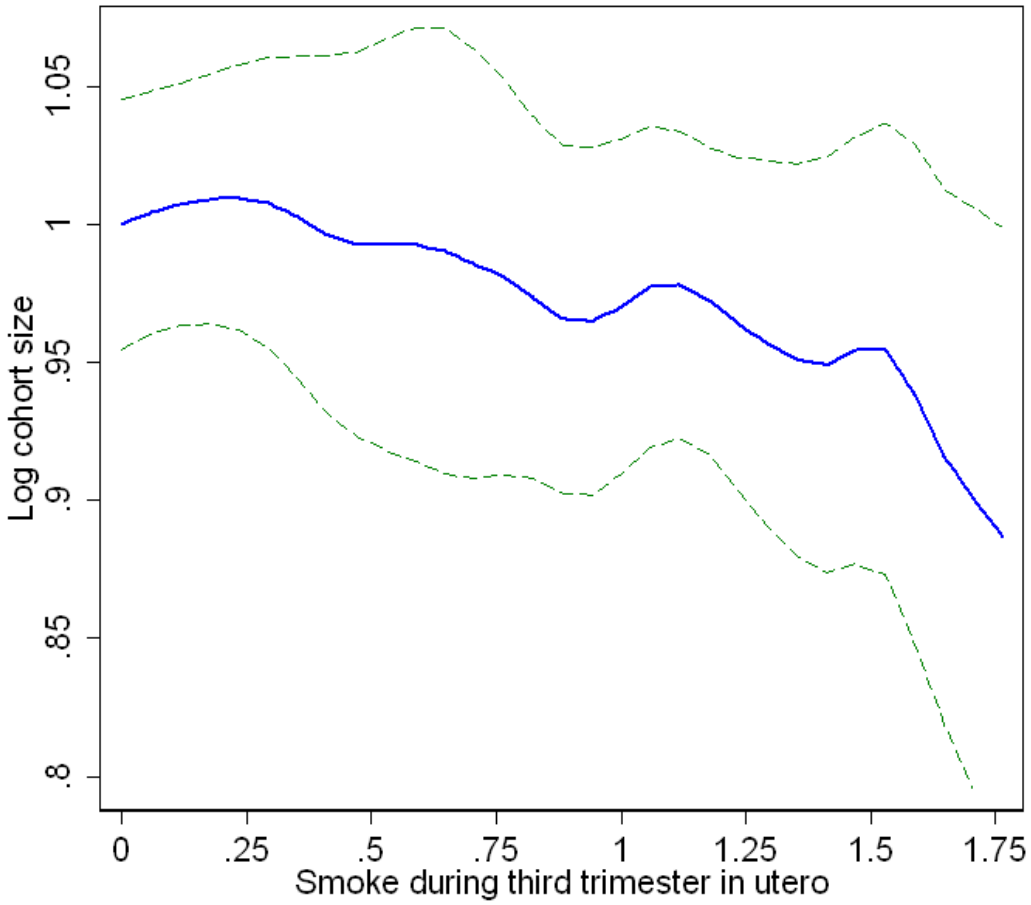


Figure 4: Kernel regression of log cohort size on prenatal exposure to pollution




The solid line is the relationship between log cohort size and pollution (*PrenatalSmoke*). The dashed lines mark the bootstrapped 95% confidence interval, with errors clustered within an island-month. The model estimated is a locally weighted non-parametric regression of log cohort size on pollution conditional on linear year and district fixed effects, following Robinson (1988). Log cohort size has been offset by a constant so that its value is 1 at an aerosol index of 0.

Figure 5: Timing of the fires and the financial crisis

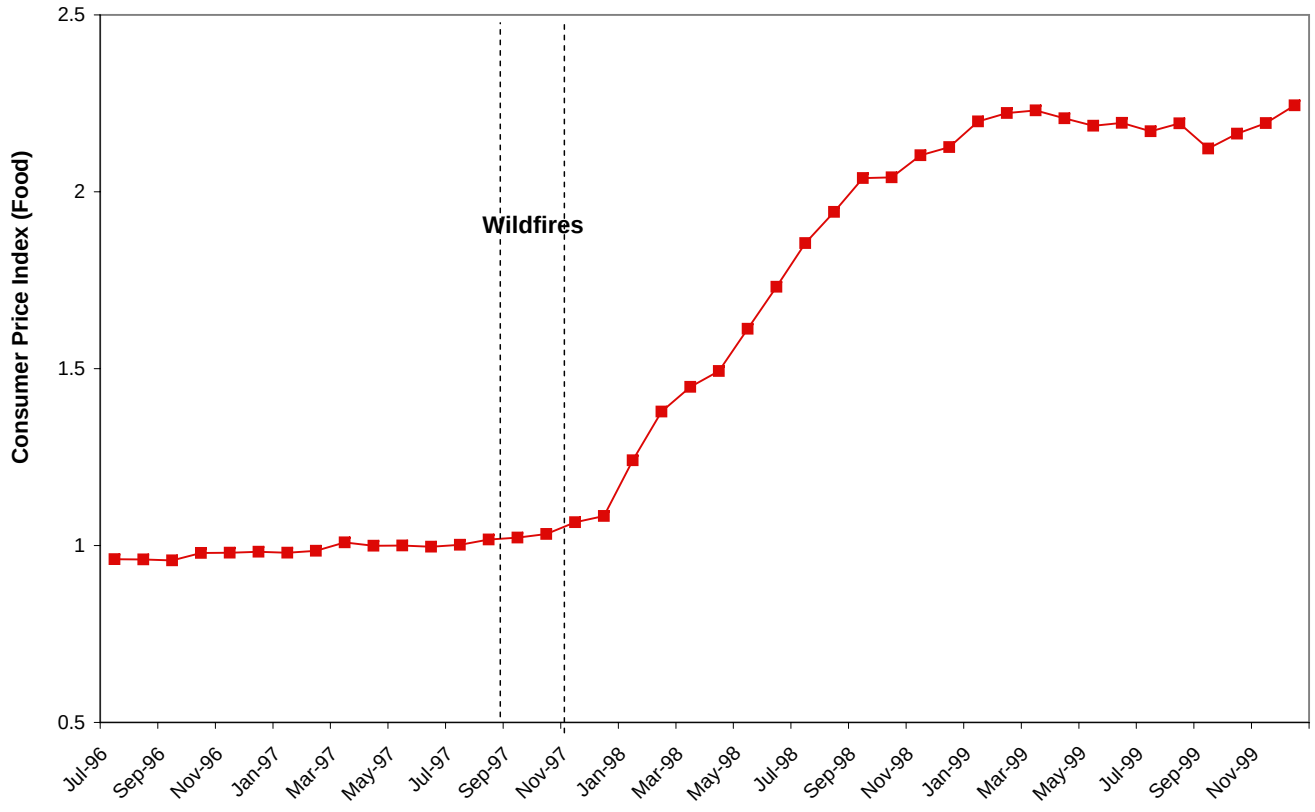


**Table 1**
**Descriptive Statistics**

| | Mean | Std. Dev. |
|---|---|---|
| Cohort size variables | | |
| Cohort size (for subdistrict-month) | 95.6 | 89.7 |
| Ln(cohort size) | 4.8 | .8 |
| Pollution variables | | |
| Smoke (median daily value for month) | .087 | .424 |
| Prenatal Smoke ($Smoke_{t-1,2,3}$) | .095 | .330 |
| Postnatal Smoke ($Smoke_{t+1,2,3}$) | .074 | .342 |
| Proportion of days with high smoke (aerosol index > .75) | .047 | .154 |
| Average smoke (daily values averaged for the month) | .120 | .445 |
| Mean of Smoke for Sept-Nov 1996 | .048 | .069 |
| Mean of Smoke for Sept-Nov 1997 | .578 | .791 |
| Mean of Prenatal Smoke for Oct 1996 - Feb 1997 | .032 | .052 |
| Mean of Prenatal Smoke for Oct 1997 - Feb 1998 | .365 | .505 |
| Other variables | | |
| Fires (any fires) | .157 | .364 |
| Intense fires (number x duration of fires >= 10 fire-days) | .026 | .157 |
| Rainfall (4 months before birth) relative to 1990-95 | .855 | .656 |
| Ln(median 1996 household food consumption) | 10.52 | .26 |
| 75th percentile | 10.71 | |
| 50th percentile | 10.49 | |
| 25th percentile | 10.33 | |
| Median HH food consumption in 1996 / Median HH food consumption in 1998 | .742 | .070 |
| National consumer price index (for food) | 1.131 | .202 |
| Urbanization | .57 | .39 |
| Wood as primary cooking fuel | .636 | .413 |
| Doctors per 1000 people | .161 | .241 |
| Maternity clinics per 1000 people | .031 | .050 |
| Educated mothers (completed junior high) | .386 | .215 |

Notes: The sample consists of 67454 subdistrict-birthmonths from December 1996 to May 1998. Sample averages are weighted by population (the number of people enumerated in the Census born in the year before the sample period), except for cohort size for which the unweighted mean is shown. Cohort size is the number of people enumerated in the 2000 Census who were born in a subdistrict in a given month. Smoke is the monthly median of the daily TOMS aerosol index which is interpolated from TOMS grid points within 100 km of the subdistrict's geographic center and weighted by the inverse distance between the grid point and subdistrict center. Prenatal and Postnatal Smoke are averages of Smoke for the three months before and after the month of birth. Fire-days is calculated from European Space Agency hot spots within 50 km of the subdistrict's center. Rainfall is measured 4 months prior to the month of birth at the nearest grid point on a 0.5 degree latitude/longitude grid. Urbanization is the subdistrict's percent of births in urban areas based on those born in 1994 to 1996 and uses an indicator in the Census of whether the respondent's locality is rural or urban. Educated mothers is the percent of infants whose mother has completed junior high and is based on matching infants to mothers in the Census. Median food consumption is a per capita measure for each household that uses data from the 1996 and 1999 SUSENAS household survey, as described in the data appendix. Consumer price index is from the Indonesian central bank. Health care variables are calculated for each subdistrict using the 1996 PODES (survey of village facilities). PODES and SUSENAS data are available for 63158 observations.

**Table 2**
**Comparison of Log Cohort Size Variable to Survey Data on Infant Mortality**

Dependent variable: Log cohort size

| | Level of an observation | | |
|---|---|---|---|
| | Quarter * province | Quarter * province * gender | Month * province |
| | (1) | (2) | (3) |
| Infant Mortality Rate | -1.34 | -.83 | -.54 |
| | (1.12) | (.65) | (.41) |
| Log Births | 1.60 | 1.11 | .83 |
| | (.29) | (.23) | (.18) |
| Male | | -.01 | |
| | | (.02) | |
| p-value of test that IMR coefficient = -1 | .78 | .79 | .27 |
| p-value of test that Log Births coefficient = 1 | .05 | .62 | .36 |
| Observations | 1248 | 2496 | 3742 |

Notes: The dependent variable is the log of the cohort size enumerated in the 2000 Census. The independent variables are from the 2002 Demographic and Health Survey. Infant Mortality Rate is the number of children who have died by June 2000 divided by all children born. Log births is all children born. There are 26 provinces in the sample, and the period covers 48 quarters from 1988 to 1999. In column 3, for 2 of the potential 3744 observations, there are no births. Standard errors allow for clustering within a province.

**Table 3**

**Relationship Between Air Pollution and Cohort Size**

Dependent variable: Log cohort size

| | Statistic used for smoke measures | | | | | | |
|---|---|---|---|---|---|---|---|
| | Median | Median | Mean | % high-smoke days | Median | Mean | % high-smoke days |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Smoke | -.0005 | | -.001 | -.010 | .001 | .018 | .035 |
| | (.006) | | (.007) | (.020) | (.009) | (.014) | (.036) |
| Prenatal Smoke ($Smoke_{t-1,2,3}$) | -.035 *** | -.032 *** | -.032 ** | -.085 ** | | | |
| | (.012) | (.011) | (.013) | (.033) | | | |
| Postnatal Smoke ($Smoke_{t+1,2,3}$) | -.014 | | -.016 * | -.042 * | | | |
| | (.009) | | (.010) | (.025) | | | |
| $Smoke_{t-1}$ | | | | | -.010 | -.028 * | -.069 * |
| | | | | | (.009) | (.016) | (.040) |
| $Smoke_{t-2}$ | | | | | -.023 *** | -.006 | -.035 |
| | | | | | (.008) | (.013) | (.038) |
| $Smoke_{t-3}$ | | | | | -.003 | -.005 | .005 |
| | | | | | (.013) | (.015) | (.030) |
| $Smoke_{t+1}$ | | | | | -.010 | -.019 | -.030 |
| | | | | | (.009) | (.014) | (.031) |
| $Smoke_{t+2}$ | | | | | -.005 | -.003 | -.034 |
| | | | | | (.008) | (.014) | (.034) |
| $Smoke_{t+3}$ | | | | | .001 | -.001 | .010 |
| | | | | | (.009) | (.012) | (.031) |
| Observations | 67454 | 67454 | 67454 | 67454 | 67454 | 67454 | 67454 |
| Subdistrict and month FEs? | Y | Y | Y | Y | Y | Y | Y |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. Observations are weighted by the number of individuals enumerated in the Census who reside in the subdistrict and were born in the year before the sample period.

**Table 4**
**Distinguishing between Mortality and Migration**

Dependent variable: Log cohort size

| | District of residence versus birthplace versus mother's 1995 residence | | |
|---|---|---|---|
| | Residence | Birthplace | Mother's 1995 residence |
| | (1) | (2) | (3) |
| Smoke | -.002 | .002 | .002 |
| | (.006) | (.006) | (.006) |
| Prenatal Smoke | -.035 *** | -.037 *** | -.038 *** |
| | (.012) | (.012) | (.012) |
| Postnatal Smoke | -.013 | -.015 | -.016 |
| | (.010) | (.010) | (.010) |
| Observations | 5829 | 5829 | 5829 |
| Fixed effects | month, district | month, district | month, district |

Notes: Each observation is a district-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. Observations are weighted by the number of individuals enumerated in the Census who reside in the district and were born in the year before the sample period.

**Table 5**
**Alternative Hypotheses**

Dependent variable: Log cohort size

| | Control for predicted fertility | Excluding September 1997 | SUSENAS and PODES subsample | Control for financial crisis | Excluding areas with fires | Control for fires | Control for rainfall |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Smoke | .001 | .0001 | .002 | .002 | .005 | .004 | .0001 |
| | (.006) | (.009) | (.006) | (.006) | (.008) | (.006) | (.006) |
| Prenatal Smoke | -.035 *** | -.035 *** | -.032 *** | -.032 *** | -.035 *** | -.032 ** | -.035 *** |
| | (.012) | (.013) | (.011) | (.011) | (.013) | (.014) | (.012) |
| Postnatal Smoke | -.014 | -.013 | -.012 | -.012 | .002 | -.005 | -.015 |
| | (.009) | (.010) | (.009) | (.009) | (.011) | (.011) | (.009) |
| Ln(Predicted Births) | .875 | | | | | | |
| | (.696) | | | | | | |
| Financial Crisis | | | | -.049 | | | |
| | | | | (.038) | | | |
| Any Fires | | | | | | -.004 | |
| | | | | | | (.010) | |
| Prenatal Any Fires | | | | | | .007 | |
| | | | | | | (.017) | |
| Postnatal Any Fires | | | | | | -.004 | |
| | | | | | | (.014) | |
| Intense Fires | | | | | | -.028 * | |
| | | | | | | (.016) | |
| Prenatal Intense Fires | | | | | | -.017 | |
| | | | | | | (.025) | |
| Postnatal Intense Fires | | | | | | -.021 | |
| | | | | | | (.029) | |
| Rainfall | | | | | | | -.004 |
| | | | | | | | (.007) |
| Observations | 67454 | 63703 | 63158 | 63158 | 60295 | 67454 | 67454 |
| Subdistrict and month FEs? | Y | Y | Y | Y | Y | Y | Y |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. Observations are weighted by the number of individuals enumerated in the Census who reside in the subdistrict and were born in the year before the sample period. Predicted Births is constructed using the fertility rate by age and the number of women of different child-bearing ages within a district, as described in the data appendix.The financial crisis variable is standardized to have a mean of 0 and standard deviation of 1 for the sample. Areas without fires are those with fewer than 80 fire-days over the entire period. Any Fires is an indicator of any fires and intense fires is an indicator of at least 10 fire-days in the month. The rainfall variable is constructed as rainfall 4 months prior to the birth month divided by the 1990-95 average for that calendar month.

**Table 6**

**Effects by Gender and Income**

Dependent variable: Log cohort size

| | By gender | By income (log consumption) of the district | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | <---------- one regression ----------> | | | |
| | | | | | Top quartile | 3rd quartile | 2nd quartile | Bottom quart. |
| | (1) | (2) | (3) | (4) | (5) | | | |
| Smoke | -.008 | -.060 *** | -.024 | -.010 | -.004 | -.011 | -.028 | .002 |
| | (.007) | (.021) | (.016) | (.007) | (.009) | (.010) | (.024) | (.045) |
| Prenatal Smoke | -.030 ** | -.158 *** | -.129 *** | -.069 *** | -.058 *** | -.076 *** | -.094 ** | -.121 ** |
| | (.012) | (.037) | (.028) | (.013) | (.018) | (.017) | (.047) | (.061) |
| Postnatal Smoke | -.019 * | -.158 *** | -.047 * | -.032 *** | -.025 | -.040 *** | -.046 | .009 |
| | (.010) | (.027) | (.024) | (.011) | (.016) | (.014) | (.032) | (.052) |
| Male | .014 *** | | | | | | | |
| | (.003) | | | | | | | |
| Smoke * Male | .016 *** | | | | | | | |
| | (.005) | | | | | | | |
| Prenatal Smoke * Male | -.009 | | | | | | | |
| | (.007) | | | | | | | |
| Postnatal Smoke * Male | .010 | | | | | | | |
| | (.006) | | | | | | | |
| Smoke * High Consum. | | .066 *** | .017 | | | | | |
| | | (.021) | (.014) | | | | | |
| Prenatal Smoke * High Consum. | | .127 *** | .072 *** | | | | | |
| | | (.038) | (.027) | | | | | |
| Postnatal Smoke * High Consum. | | .161 *** | .017 | | | | | |
| | | (.026) | (.014) | | | | | |
| Observations | 134734 | 63158 | 63158 | 63158 | <---------- | 63158 | ----------> | |
| Fixed effects included | subdistrict, month | subdistrict, month | subdistrict, month * high cons. | subdistrict, month * high cons. | subdistrict, month*quartile of log consumption | | | |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. High consum. is an indicator that equals 1 if the district's median log food consumption is above the sample median. Observations are weighted by the number of individuals enumerated in the Census who reside in the subdistrict and were born in the year before the sample period.

**Table 7**

**Effects By Urbanization, Wood Fuel Use, and Health Care Sector**

Dependent variable: Log cohort size

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Prenatal Smoke | -.121 *** | .015 | -.115 *** | -.113 *** | -.007 |
| | (.028) | (.032) | (.027) | (.028) | (.025) |
| Prenatal Smoke * Urbanization | -.013 | | | | |
| | (.013) | | | | |
| Prenatal Smoke * Wood Fuel Use | | -.155 *** | | | -.120 *** |
| | | (.036) | | | (.026) |
| Prenatal Smoke * Maternity Clinic | | | .030 *** | | .011 ** |
| | | | (.009) | | (.005) |
| Prenatal Smoke * Doctors | | | | .048 *** | .016 |
| | | | | (.015) | (.013) |
| Prenatal Smoke * High Consum. | .071 *** | .048 * | .058 ** | .052 ** | .044 * |
| | (.027) | (.025) | (.025) | (.025) | (.025) |
| Observations | 63158 | 63158 | 63158 | 63158 | 63158 |
| Subdistrict and month FEs? | Y | Y | Y | Y | Y |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. All regressions also include Smoke and Postnatal Smoke and their interactions with the relevant variables for each column. Urbanization is the proportion of the population in urban localities and is based on 1994 to 1996 birth cohorts. Wood fuel use is an approximate measure of the proportion of people in the subdistrict who cook with wood fuel rather than kerosene and gas. Health variables are normalized to be mean 0, standard deviation 1 for the sample. High consum. is an indicator that equals 1 if the district's median log food consumption is above the sample median. Observations are weighted by the number of individuals enumerated in the Census who reside in the subdistrict and were born in the year before the sample period.

**Table 8**

**Effects by Mother's Education**

Dependent variable: Log cohort size

| | Measure of mother's education | |
|---|---|---|
| | Individual-specific | Subdistrict average |
| | (1) | (2) |
| Smoke | -.012 | -.013 |
| | (.022) | (.017) |
| Prenatal Smoke | -.075 *** | -.113 *** |
| | (.025) | (.029) |
| Postnatal Smoke | -.065 ** | -.044 |
| | (.027) | (.028) |
| Smoke * Educated Mother | .021 | |
| | (.041) | |
| Prenatal Smoke * Educated Mother | .075 | |
| | (.048) | |
| Postnatal Smoke * Educated Mother | .098 * | |
| | (.052) | |
| Educated Mother (junior high +) | -.513 *** | |
| | (.089) | |
| Smoke * % Educated Mothers | | -.009 |
| | | (.020) |
| Prenatal Smoke * % Educated Mothers | | -.034 |
| | | (.028) |
| Postnatal Smoke * % Educated Mothers | | .053 * |
| | | (.034) |
| Smoke * High Consumption | | .018 |
| | | (.014) |
| Prenatal Smoke * High Consumption | | .099 *** |
| | | (.026) |
| Postnatal Smoke * High Consumption | | .011 |
| | | (.034) |
| Observations | 134908 | 63158 |
| Fixed effects included | subdistrict, month * educated mother | subdistrict, month * high consum., month * % educated mothers |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. In column 1, each cell is a subdistrict-month-education category. Educated mothers are defined as those who have completed junior high. In column 2, each cell is a subdistrict-month, and % educated mothers is the subdistrict average over the sample period. When interacted with month fixed effects, % educated is standardized to mean 0. High consumption is an indicator that equals 1 if the district's median log food consumption is above the sample median.

**Table A1**
**Instrumental Variables Estimation**

| | First stage | IV |
|---|---|---|
| Dependent variable | Prenatal Smoke | Log cohort size |
| | (1) | (2) |
| (Sumatra or Kalimantan) * (Oct 97 to Jan 98) | .724 *** (.094) | |
| Prenatal Smoke | | -.040 ** (.016) |
| Observations | 67454 | 67454 |
| Fixed effects | month, subdistrict | month, subdistrict |
| F-statistic for instrument | 59.0 | n/a |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05. Observations are weighted by the number of individuals enumerated in the Census who reside in the subdistrict and were born in the year before the sample period.

**Table A2**
**Different Sample Periods**

Dependent variable: Log cohort size

| | Shorter periods | | | Balanced calendar months | | Longer period |
|---|---|---|---|---|---|---|
| | 3/97 - 5/98 | 12/96 - 2/98 | 3/97 - 2/98 | 11/96 - 2/97 & 10/97 - 2/98 | 11/96 - 2/97 & 11/97 - 2/98 | 12/96 - 5/98 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Smoke | -.003 | -.004 | -.009 * | -.005 | -.021 | -.001 |
| | (.005) | (.005) | (.005) | (.009) | (.042) | (.006) |
| Prenatal Smoke | -.043 *** | -.036 *** | -.049 *** | -.026 * | -.030 ** | -.026 ** |
| | (.012) | (.012) | (.012) | (.014) | (.014) | (.012) |
| Postnatal Smoke | -.023 *** | -.012 | -.026 *** | .025 | .038 | -.006 |
| | (.009) | (.010) | (.008) | (.030) | (.032) | (.009) |
| Observations | 56220 | 56201 | 44967 | 33684 | 29933 | 78703 |
| Subdistrict & month FEs? | Y | Y | Y | Y | Y | Y |

Notes: Each observation is a subdistrict-month. Standard errors, in parentheses below the coefficients, allow for clustering at the island-month level. *** indicates p<.01; ** indicates p<.05, * indicates p<.10. Observations are weighted by the number of individuals enumerated in the Census who reside in the subdistrict and were born in the year before the baseline sample period.