Air Quality and Infant Mortality During Indonesia's Massive Wildfires in 1997

Seema Jayachandran*

March 2005

Abstract

Air quality was extremely poor in Indonesia in late 1997 due to smoke from massive wildfires (which had been set intentionally to clear agricultural land but ran rampant). This paper examines the impact this episode of air pollution had on infant and fetal mortality. Infant and fetal death are inferred from "missing children" in the 2000 Indonesian Census; by analyzing subdistrict-birthyear-birthmonth cohorts, one can exploit the detailed timing and spatial patterns of the pollution. Exposure to higher levels of airborne smoke, especially during the last trimester in utero, is associated with a substantial decline in the size of the surviving cohort. The fire-induced pollution caused a 1.0% decrease in cohort size (or over 16,400 missing children), averaged across Indonesia for the five-month period of high exposure. Indonesia's baseline under–2 mortality rate was about 6%, so if pollution mainly caused infant rather than fetal deaths, the estimates would represent up to a 17% effect. Prenatal exposure seems to have a larger effect on boys, while postnatal exposure has a larger impact on girls, suggesting different physiological vulnerabilities and possibly gender discrimination. In addition, pollution has much larger effects in poorer areas.

*Robert Wood Johnson Foundation Scholar in Health Policy Research, UC Berkeley; Assistant Professor, Dept. of Economics, UCLA. e-mail: jaya@ucla.edu. I thank Michael Kremer, Ben Olken and seminar participants at UC Berkeley for helpful comments; Kok-Hoe Chan and Daniel Chen for data; and Nadeem Karmali and Rika Christanto for excellent research assistance.

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, palm oil producers, or small farmers clearing land to plant new crops or trees.[1] Because of the dry, windy conditions and delayed rainy season 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 poor air quality. The daily level of airborne smoke for different locations in Indonesia is available from satellite measurements. The outcome, infant and fetal mortality, is inferred from "missing children" in the 2000 Census. This approach overcomes the problem that mortality records are unavailable for Indonesia and survey data on infant mortality have small sample sizes. The paper finds that higher levels of pollution are associated with substantial declines in the size of the surviving cohort, and that it is exposure to pollution during the last trimester in utero that has the largest impact on survival. The fire-induced increase in air pollution caused 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 about 6%, so if the effect of pollution operated mainly through infant deaths, this would represent up to a 17% effect.

The estimates imply that over 16,400 infant and fetal deaths are attributable to the fires. Cost estimates of the fires have focused mainly on destroyed timber, low productivity during the pollution episode, lost tourism, and the like. These costs are estimated at $2 to 3 billion (Tacconi 2003). The health costs of the fires are likely much larger: Even if each human life were valued at the modest amount of $100,000, the infant mortality costs alone

[1]The Indonesian Minister of Forestry estimated that "[commercial] plantations caused some 80% of the forest fires," and that small farmers practicing slash-and-burn harvesting caused the remainder (*Straits Times*, September 3, 1997). Rabindran (2001), using satellite data on land use, finds that the 1997 incidence of fires on commercial 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.

would be close to $2 billion.

There is also a striking amount of heterogeneity in the impact of pollution. The effect size is much larger in poorer areas. Pollution has twice the effect in districts whose consumption level is above the sample median compared to those below the median. Among possible explanations, individuals in poorer areas could be more susceptible because they have lower baseline health, or they could have less access to health care. There is suggestive evidence in support of these hypotheses. Pollution causes more mortality in areas with fewer medical facilities and personnel. In addition, people who also are exposed on a daily basis to indoor air pollution might suffer more acute health effects from the added pollution from the wildfires, and the estimated effects are indeed larger in areas where more people use wood-burning stoves. A third hypothesis is that it is urbanization that matters because in rural areas people are more likely to work outdoors, or there are fewer and less effective information campaigns advising the public on precautions to take, for example. However, the heterogeneity by income is in fact not explained by the rural versus urban distinction.

Pollution also has differential effects by gender. Exposure during the month of birth and months after birth has a larger impact on girls, possibly reflecting gender discrimination. Parents might make more effort to shield their sons from pollution or seek medical care for them. Prenatal pollution, on the other hand, seems to be somewhat more detrimental to males, consistent with previous findings that male fetuses are physiologically less robust than female fetuses.

The 1997 episode was extraordinary, but widespread fires are not a rarity in Indonesia and in other countries, particularly in Southeast Asia, that use fire to clear land. In Indonesia, illegal logging, which was made possible by lax enforcement of logging regulations, was a contributing factor in the fires. The findings of this paper highlight a large cost of such policies.

In addition, using the sharp timing of the Indonesian fires, one can isolate and identify the effects of pollution to answer broader questions about pollution and health. First, individuals in developing countries are exposed to air pollution from several sources, and

the estimates in this paper suggest that even at moderate levels of pollution, reductions in pollution would save many infant lives. The wildfires are comparable in particular to wood-burning stoves which produce a similar level and mix of pollutants when used indoors, a practice that is common in developing countries. Second, the paper finds that exposure in utero is especially important. The previous literature has hinted at but not firmly established this result. One policy implication may be that the priority of public health efforts in response to pollution should be to target pregnant women. Third, people in poorer areas, and perhaps infant girls as well, are particularly vulnerable to pollution. The glass-half-full interpretation is that as communities make strides in economic development and in the treatment of girls, the mortality effects of pollution could become less severe. The glass-half-empty view is that, in the meantime, the health effects of pollution in developing countries are regressive in addition to being large. While caution is warranted when extrapolating from air pollution caused by the fires to other settings, the findings of this paper likely have broader applicability.

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

A growing body of work examines the relationship between air quality and infant mortality, primarily in the United States. Chay and Greenstone (2003b) use geographic variation in the extent to which the 1980–81 recession lowered pollution in the United States and find that better air quality reduces infant deaths. Chay and Greenstone (2003a) find that air pollution abatement after passage of the Clean Air Act of 1970 led to a decline in infant deaths.

Currie and Neidell (2004) are able to examine how the timing of exposure to pollution affects health by using California birth and death records. Using within-zipcode variation over the 1990's, they find evidence that exposure to pollution during the month of birth causes infant mortality but find no effects on low birthweight, premature birth, or fetal death. Like Currie and Neidell (2004), this paper looks closely at the timing of exposure to pollution.[2]

In addition, there have been studies on primarily the adult health effects of Indonesia's 1997 fires. Sastry (2002) examines the impact of the fires on mortality in nearby Malaysia and finds increased mortality for older adults on the day following a high-pollution day. Frankenberg, McKee, and Thomas (2004) study adult health outcomes with a differences-in-differences approach that compares outcomes in 1993 and 1997 for areas with high versus low exposure to the 1997 fires. They find that pollution from the fires reduced the ability to perform strenuous tasks and other measures of health.

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 (μm) and especially less than 2.5 μm in diameter are considered the most harmful to health because they are small enough to be inhaled and transported deep into the lungs. For biomass smoke, the modal size of particles is between 0.2 and 0.4 μm, and 80 to 95% of particles are smaller than 2.5 μm (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 either because pollution inhaled by the mother and absorbed into her bloodstream interferes with her health which in turn disrupts fetal nutrition and

[2]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 smog episode.

oxygen flow, or because toxicants cross the placenta. Several studies find an association between air pollution and retardation of fetal growth or shorter gestation period, both of which are associated with lower birthweight. Recently, for example, Berkowitz et al. (2003) examined New York City women who were pregnant during the World Trade Center disaster on September 11, 2001, and found that intrauterine growth retardation was significantly more prevalent among women who were near the WTC during the event compared to those who were not.[3] In addition, there is some evidence on the biological mechanisms behind these pregnancy outcomes. One type of toxicant prevalent in biomass smoke is polycyclic aromatic hydrocarbons (PAH). Air pollution has been associated with a greater prevalence of PAH-DNA adducts on the placenta (PAH bound to DNA), which are a sign of damage to the genome. PAH-DNA adduct levels, in turn, have been found to be correlated with low birth weight and small head size (Perera et al. 1998, Topinka et al. 1997). In addition to retarding fetal growth, in utero exposure to PAHs and other pollutants has been linked to increased risk of infant leukemia (Alexander et al. 2001).

2.2 Description of the Indonesian fires

Even the dry season in Indonesia, which occurs from July to September, is typically wet. Monthly dry season rainfall averages about 10 centimeters (cm) in both Sumatra and Kalimantan, the Indonesian part of Borneo. (See Figure 1 for a map of Indonesia.) However, the dry season in 1997 was particularly dry. Figure 2 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 in July, August, and September was lower than usual, and the rainy season was delayed until November. Similarly, 30 cm of rainfall were recorded in East Kalimantan for the 12 months beginning in April 1997 compared to the typical amount of 270 cm.

Fires are commonly used in Indonesia to clear land for cultivation, and the dry season is

[3]International evidence on adverse pregnancy outcomes associated with particulate matter includes Dejmek et al. 1999 on the Czech Republic and Wang et al. 1997 on China, among others.

considered an opportune time to set fires because the vegetation burns quickly. Industrial farmers clear land for replanting forests as palm trees, and many small farmers rotate among several plots using swiddening techniques in which land is cleared, cultivated, and then left fallow for several years. 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 area, they often left behind waste wood, even though it had economic value as fertilizer or wood chips. This left-behind wood made the forest more susceptible to fast-spreading fires.[5]

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 the El Niño conditions affected the rest of Indonesia as well. What set Sumatra and Kalimantan apart is that Indonesia's forests and other types of land most susceptible to rampant fires are mainly in these areas. The majority of

[4]The timber industry arose in 1967 when the new Suharto government began encouraging foreign investment and passed the Basic Forestry Law which gave the central government authority to grant logging concessions. The palm oil industry also began in the late 1960's and grew rapidly through the mid-1990's. In 1996 forest products accounted for 10% of Indonesia's gross domestic product, and Indonesia supplied about 30% of the world palm oil market (Barber and Schweithhelm 2000, Ross 2001).

[5]See Barber and Schweithhelm (2000) for a comprehensive account of how commercial practices may have contributed to the fires.

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 tracked the location of the fires, though due to wind patterns, not perfectly. Figure 3 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. There was relatively little smoke over Java while the northern half of Sumatra was heavily affected, though there were few fires in either area.

The level of pollution reached in Indonesia was equivalent to smoking several packs of cigarettes a day. A common measure of particulate matter is PM_{10}, the concentration of particles less than 10 μm 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. In September 1997, PM_{10} levels reportedly reached as high as 3000 $\mu g/m^3$ in Sumatra and Kalimantan. In the hardest hit areas, pollution surpassed 1000 $\mu g/m^3$ on several days and exceeded the EPA standard of 150 $\mu g/m^3$ for long periods (Ostermann and Brauer 2001, Heil and Goldmammer 2001). The levels of pollution caused by the wildfires are comparable to levels caused by indoor use of wood-burning stoves. The daily average pollution 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).

A second common measure of pollution is the Pollution Standards Index (PSI) which is a composite measure of carbon monoxide, nitrogen dioxide, sodium dioxide, ozone, and PM_{10}. The PSI scale runs from 0 to nominally 500, with a level above 400 considered life-threatening to ill and elderly people. PSI levels of over 500 were recorded in Indonesia on several days during September to November 1997. One reason the Indonesian fires produced so much pollution is that many of the fires were peat fires which produce large amounts of smoke.[6]

[6]Less important for immediate health effects but of possibly long-term environmental consequences, the fires released a large amount of carbon into the atmosphere. The fires were one cause of the unusually large

3 Empirical Strategy and Data

3.1 Empirical model and outcome measure

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:

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

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 β_1 is negative, or that exposure to smoke 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 sample has on average 1 birth and 0.05 recorded child deaths per district-month for the affected cohorts. The Census population module, a long-form survey administered to a sample of 2000 Census respondents, records approximately 7 births and 0.3 deaths per district-month. Survey data on infant mortality are generally considered poor in quality, as well.

Thus, the approach I take is to infer fetal and infant mortality by measuring "missing children."[7] 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

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

jump in atmospheric carbon dioxide levels between 1997 to 1998 (Page et al. 2002).

[7]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.

The dependent variable, $ln(CohortSize)_{jt}$, is the natural logarithm of the number 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 twelve months prior to the sample period).

The main advantage of inferring deaths by counting survivors is that one can used data 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 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 one and a half. For a younger cohort born in May 1998, the outcome is survival until age two, and for an older cohort born in December 1996, the outcome is survival until age two and a half.[8] The inclusion of birthyear-birthmonth (hereafter, month) fixed effects in the regression will control for any average differences in survival by cohort.[9]

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

[8] As shorthand I describe deaths of children in the sample as infant mortality even though they could occur as late as age two and a half. The common definition of infant mortality is deaths before age one.

[9] One advantage of observing survival over a year after the due date is that if the deaths occur as stillbirths or neonatal deaths, the estimates are less likely to reflect simply short-term "harvesting."

that the results are an artefact 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 just after birth that affects mortality. The motivation for this model are findings from previous research that exposure 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 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 the fires or afterwards), then cohort size would be smaller in affected areas. Fortunately, one can partially allay the 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.[10] There are on average 96 surviving

[10]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

children per month in the sample. Note that 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

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

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.[11] As mentioned above, very few births per subdistrict-month are sampled in the DHS, so I aggregate to provinces and quarters. In addition, I use a longer panel from 1988 to 1999. 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, but the IMR variable should not be affected by this potential 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

migrants in 1999 (in response to ethnic violence against them in Kalimantan), and also the province of Aceh 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.

[11] The comparison would be better if the DHS also recorded pregnancies that ended in fetal deaths because some of the missing children in the Census are not among the live births measured by the DHS.

bias from measurement error when a smaller and hence noisier cell size is 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 since it is unlikely that there are large fluctuations in fertility that coincide with the air pollution from the fires both spatially and temporally. Even area-specific trends could not explain the patterns since the sample includes time periods both before and after the fires; any omitted fertility shift 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 finds that fluctuations in predicted fertility do not seem to be a confounding factor.

3.3 Pollution measure

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 aerosol optical depth, or the amount of light that microscopic airborne particles absorb or reflect. The TOMS aerosol index has been found to match quite closely data collected by ground-based pollution monitors (Hsu et al. 1999). Ground-based data on suspended particulates 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 value is more smoke.[12]

The TOMS data set contains daily aerosol measures (which are constructed from obser-

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

vations 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 the mean of the daily values and the number of days that exceed a (somewhat arbitrary) threshold value of 0.3.

Whereas there are over 3700 subdistricts in the data, there are only 226 unique pollution grid points used. Interpolation adds variation at a finer grain spatially, but uncorrected standard errors would still 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 the pollution level in the month of birth ($Smoke_{jt}$) as well as lags of $Smoke_{jt}$ which measure exposure 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 0.3 on 17% 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 level of the *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 during the fires. Figure 4 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 late 1997, also experienced another episode of smoke in early 1998 after the rainy season ended. Figure 5 plots the corresponding outcome data, that is, the aggregate cohort size (seasonally adjusted) for Kalimantan, Sumatra, and the rest of Indonesia. Cohort size is noisy, but the raw data foreshadow the regression results. In Sumatra and especially Kalimantan, there is a dip down among cohorts born just after the fires, suggesting that prenatal exposure to the smoke led to mortality. The regression analysis presented below, which uses more spatial variation in pollution and controls for month and subdistrict fixed effects, finds similar results.

3.4 Other variables

Several other variables are used in the analysis 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.

In addition, 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 infrastructure and other 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.

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 number of surviving children in a cohort. *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 but 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.[13] Columns 3 and 4 consider alternative monthly pollution

[13]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

measures, first, the mean rather than median of the daily pollution values and, second, the proportion of days with high pollution (aerosol index above 0.3). The mean pollution gives nearly identical results as the median value, with exposure in the quarter after birth having a negative impact on cohort size that is marginally statistically 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 0.9%.

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 during the fires. Multiplying that gap by the coefficient of -0.035 suggests 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. Indonesia's baseline under-2 mortality rate was roughly 60 per 1000 live births at this time, based on government statistics.[14] If the effects of pollution were due exclusively to infant deaths, the estimates would represent a 17% effect; if the effects were due in equal parts to infant and fetal deaths, the coefficient would imply an 8% effect.

The welfare implications of mortality caused by the pollution depend on the counterfactual of how long the children would have lived. One way of gauging whether the results could be due to "harvesting" is to compare the baseline under-2 and under-5 mortality rates for Indonesia. A little more than 1% of children who survive until age 2 die by age 5. To attribute the 1% effect to harvesting, essentially all deaths between age 2 and 5 would

variation in pollution attributable to the fires, is -0.040.

[14]See Table A2 for infant mortality statistics. I assume that half of the additional deaths between age 1 and age 5 occur before age two.

have to have been moved forward to the time of the fires. Moreover, by most standards decreasing a child's life by two to five years would be a significant welfare loss.

Figure 7 shows the nonparametric relationship between third-trimester exposure and cohort size. The effect of pollution appears to be linear for the most part. There is some evidence of a steeper relationship at high levels of pollution, though the data are sparse in this region and the nonlinearities are imprecisely estimated.

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*), 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 remain the same, but, interestingly, the pattern for prenatal exposure is a bit different for each of the pollution measures. Using the mean level (column 6), exposure in the month immediately preceding month has the strongest negative relationship with cohort size, while using high-smoke days (column 7), exposure three months before birth has the strongest effect. One interpretation of these patterns is that at different points during gestation, fetuses are more vulnerable to sustained versus short-term, intense smoke. A likelier interpretation is that there is not enough precision to determine at this finer grain how the timing of exposure affects survival. Thus, for the rest of the analysis, I focus on the three-month measures of prenatal and postnatal exposure.

4.2 Effect of smoke on mortality versus alternative hypotheses

The results in Table 3 suggest that exposure to smoke in utero caused infant and fetal deaths, but there are other possible explanations for the results. This section considers some alternative explanations.

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.

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, and migration could still explain the results. 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. Thus, to test for migration that occurs *before* birth, I match infants to their mothers 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, offering reassurance that between-district migration either before after birth is not an important confounding factor.

To gauge whether within-district migration could be driving the results, column 4 estimates a model with district-month fixed effects. The coefficient for $PrenatalSmoke$ is imprecise but the point estimate of -0.013 is smaller than in the main specification that also used between-district variation (Table 3, column 1). In sum, migration cannot easily account for the negative relationship between exposure to pollution and cohort size.

Fertility

The empirical approach interprets decreases in $Ln(CohortSize)$ as increases in infant deaths, but the surviving cohort also becomes smaller when births decline. It seems unlikely that conceptions decreased nine months before the fires with a spatial pattern matching the pollution from the fires, 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 through May 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, but because the measure is noisy, the estimate may suffer from downward bias. The actual coefficient on predicted births is 0.35 and statistically indistinguishable from 1. Moreover, 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.[15]

[15]Table A3 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.

Preterm births

Another hypothesis is that the missing children are not deaths, but rather shifts in the duration of pregnancies. In particular, exposure to pollution may have induced preterm births which are 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 higher since it incorporates the pollution in September. If infants due in October were instead born in September, then births would have shifted from a high-$PrenatalSmoke$ month to a low-$PrenatalSmoke$ month, generating a negative relationship between $PrenatalSmoke$ and cohort size.[16] 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 $Smoke$ would become more negative and the coefficient on $PrenatalSmoke$ would become less 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.[17]

Financial crisis

The Indonesian financial crisis began shortly after the 1997 episode of air pollution, as shown in Figure 8, 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,

[16] 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.

[17] Table A3 restricts or expands the sample to several other time periods, and the results are robust to the different sample definitions. One noteworthy finding is that the estimated effect of $PrenatalSmoke$ is smaller when the window is extended beyond 8 months after pregnancy, suggesting that the fires may have reduced fertility as well.

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 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 crisis 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). Column 5 instead uses the average of the crisis measure for the three months following the month of birth. Since the crisis accelerated a few months after the fires, this measure has more variation during the time period of interest. The estimated effect of $PrenatalSmoke$ on cohort size remains unchanged.

Effect of pollution versus reduced-form effect of the fires

Another interpretation of the results is that they represent mortality effects of the fires rather than 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. One way to separate the effect of pollution from other effects of the fires is to compare

Kalimantan and Sumatra, where the vast majority of fires occurred, to other parts of Indonesia that experienced few fires but nevertheless were exposed to smoke. In Table 5, column 6, the pollution variables are interacted with indicator variables for Kalimantan and Sumatra. If it is other effects of the fires and not pollution that matters, the main effect for pollution should be zero. The estimates are imprecise, but the point estimate for the main effect of *PrenatalSmoke* is -0.022, about two thirds the magnitude seen earlier. The *PrenatalSmoke* interaction term for Kalimantan is negative and of comparable size, while the interaction terms for Sumatra are close to zero. While this test has limited power to separately identify the effects in the regions with and without fires, it appears that there is a negative relationship pattern between pollution and survival throughout Indonesia, with perhaps larger effects in Kalimantan (consistent with the nonparametric relationship seen in Figure 7).

4.3 Effects by gender and income

Gender

This section examines how the mortality effects of pollution vary for different 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 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 imprecisely estimated but is consistent with findings in the literature that male fetuses are less physiologically robust than female fetuses (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 as well. Girls could be more susceptible to respiratory infection and other postnatal complications from pollution. The negative interaction effects for girls are also consistent with gender discrimination. If parents are more likely to take steps to

shield their male children from pollution or to seek medical treatment for acute respiratory infection, for example, then one might expect the effects of pollution to be stronger for girls.[18] One signature of discrimination would be if the gender gap were larger for Muslim children, a test that is merely suggestive but is motivated by an extensive literature that documents excess female mortality in the Middle East and often attributes it to Islamic traditions.[19] In results not reported, the point estimates suggest that the gap is slightly larger among Muslims, but the results are very imprecise and somewhat sensitive to the specification used.

The smaller postnatal effects for boys also provide additional suggestive evidence against migration as the explanation for the relationship between pollution and cohort size. Under the assumption that parents favor boys, one would expect parents of boys to be more likely to move away from hard-hit areas. Contrary to this prediction, the correlation of postnatal exposure to pollution and cohort size is weaker for boys.

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 which makes them more susceptible to pollution, have less access to health care, or have effectively more exposure to pollution because they spend more time working outdoors or perform more strenuous tasks. 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

[18] If the health effects of prenatal exposure assert their symptoms after birth, one would expect boys to also be less affected by prenatal exposure. This effect on survival would offset any physiological disadvantage of male fetuses.

[19] See Yount (2001) for a review of the literature on gender discrimination and excess female mortality in Islamic countries. Besides Islam, the other common religions in Indonesia are Christianity, Hinduism, and Buddhism. The gender gap for education in Indonesia, calculated from the 2000 Census, is larger among Muslims. It is also worth noting that the main effect for male in column 1, which measures the male-to-female sex ratio averaged across one- to two-year-olds, is 1.014 which is lower than most countries.

half of the distribution are large and positive. The model estimated in column 2 appears to be misspecified, however. Notice that 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 this apparent paradox is that the month effects vary significantly with income. As has been documented in the demography literature, seasonality in fertility and infant mortality 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 and is the preferred specification. 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. The 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 consumption level also 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 now -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 coefficient on $PrenatalSmoke$ become successively more negative moving from the top to the bottom quartile. The coefficients for the other smoke variables are imprecise, especially for the bottom two quartiles, and the point estimates do not monotonically decline with consumption. The above- and below-median consumption

levels, as opposed to a linear interaction term, seems to be the best way to parsimoniously characterize the heterogeneity by income.

Effects by urbanization, wood-stove use, and health care facilities

There are several possible reasons for the income gradient in the effects of pollution, and this subsection tests some hypotheses. It will be difficult to draw strong conclusions both because the measures used below could be correlated with omitted variables and because data are available to test only a limited number of hypotheses.

One hypothesis is that urban areas experienced smaller effects from the fires than rural areas, and it is this fact that generates the heterogeneity by income. Urbanization itself 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, and there is less outdoor employment. On the other hand, pollution from the fires may have been particularly noxious in urban areas where it mixed with industrial pollution from cars and factories. Column 1 of Table 7 interacts the pollution measures with the proportion of the population that is urban (based on those born in the two years before the sample period). Only the coefficients for $PrenatalSmoke$ and its interaction terms are reported, but $Smoke$, $PostnatalSmoke$ and their interactions are also included as well as interactions of the pollution variables with an indicator for above-median consumption. The effects of pollution do not vary by urbanization level, suggesting that the offsetting effects described above may have cancelled each other out. 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 women who work in agriculture experience larger effects.

Next I test whether the effects depend on the prevalence of wood-burning stoves. Indoor air pollution from wood-burning stoves has a similar composition and similar health effects as pollution from the wildfires. If the health impact of pollution is convex in exposure, those who have daily exposure to indoor 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 what proportion of people in the subdistrict 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 negative effects from the wildfire smoke. A 20% 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.

Finally, I examine whether the effects vary with the availability of health care workers and facilities in the area. A better health care system could lead to better baseline health as well as better medical care for illness caused by the pollution. Most health care is government-provided in Indonesia, but private care also plays a role. For example, maternity clinics, which provide prenatal and postnatal care and sometimes inpatient childbirth services, have expanded recently and an increasing share are privately run.

One empirical limitation is that the different health care measures are correlated with each other which makes it difficult to separately identify the effects of each type of health care input, so I consider a subset of measures that seem most likely to have an impact. Table 7, columns 3 to 5, present the results when interactions of the pollution measures with maternity clinics, doctors, and midwives 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 findings of Frankenberg (1995). She examines within-village changes in the health sector between 1983 and 1986 and, using household survey data, finds that infant mortality decreases when a village acquires more

maternity clinics and doctors. The Indonesian government trained and deployed thousands of midwives during the 1990's, but, as seen in column 5, the number of midwives in an area is not associated with smaller effects of pollution. This finding does not rule out that midwives had a positive effect, since midwives may have been deployed to places that were poorer or otherwise more vulnerable to pollution (though the midwife variable is nearly uncorrelated with log consumption).

In column 6, the interactions with wood fuel use, doctors, and maternity clinics are estimated in a single regression. The effects of in utero exposure to the wildfires 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 about these results is that the interaction terms do not fully explain the differential effects by income. There remains an unexplained positive interaction term for richer areas. Better measures of indoor air pollution or access to health care might explain more of the heterogeneity by income level. There are also potential channels not tested here. Mothers and newborns in poorer areas might have poorer nutrition and hence lower baseline health. Behavioral responses might differ by income with those in richer areas being more likely to use surgical masks, avoid of strenuous activity, or temporarily evacuate to less affected areas, for example.[20] But regardless of the underlying reason, the differential effects by income suggest that the mortality costs of pollution—and of the policies that contributed to the 1997 Indonesian fires—were not only large but also inequitable, with a disproportionate burden falling on 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.

[20]Kunii et al. 2002 surveyed 532 people during the fires and found that use of surgical masks among adults was associated with less severe respiratory problems from the smoke.

This paper exploits the sharp timing of the pollution and spatial variation across Indonesia to identify these effects. The paper also finds clear patterns in how the timing of exposure matters: in utero exposure to pollution has the largest effect on survival. At levels of pollution that are typical both indoors and outdoors in many poor countries, pollution has a sizeable effect on mortality.

Questions in developing countries sometimes go unstudied because the data are simply not available. Research on the infant mortality effects of air pollution in the United States makes use of 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 introduces potential problems, the paper is able to show that migration, changes in gestation period, and other potential concerns do not seem to be driving the results. Second, in lieu of ground-based pollution data, pollution data from a satellite-borne spectrometer are used. Because of the satellite's global coverage, the data cover even underdeveloped and remote areas.

The paper highlights two important ways in which environmental issues and economic development interact. First, corruption, which is prevalent in Indonesia as in many developing countries, was an important driving force behind the catastrophic fires. The Indonesian government turned a blind eye when large firms set fires in violation of the law and was loathe to crack down even at the height of the fires. The Minister of Forestry was a lone voice trying to hold companies accountable, and in late September he named 176 firms that were suspected of illegally setting fires. However, the government never followed through and in virtually no cases were timber licenses revoked or firms otherwise punished. In fact, in an act brazen by even Suharto's formidable standards of crony capitalism, in early 1998 he appointed his close associate Bob Hasan, a timber magnate who was outspoken about blaming small farmers for the fires and absolving commercial interests, to take over as Minister of Forestry. After the financial crisis, the IMF and World Bank required that Indonesia reform its forest policies as a condition of its bailout, and Indonesia has since

introduced some competitive bidding for timber concession rights and other improvements. However, its environmental policies still have many problems, and Indonesia remains prone to widespread fires.

The findings of the paper highlight a second link between the environment and economic development: the health burden from pollutants falls disproportionately on the poor. The estimated effect size is strikingly larger in poorer areas compared to richer areas. There is some suggestive evidence that this 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 seems to be the limited access to health care in poorer areas. For the most part, though, the underlying cause of the differential effects by income is an open question. Another provocative pattern is that postnatal exposure to pollutants has a larger impact on girls, which could be due to greater physiological vulnerability but also raises the possibility of gender discrimination. Why the health effects of pollution differ between rich and poor and between boys and girls is one area to pursue to better understand how environmental degradation may create unique concerns and challenges in developing countries.

Data Appendix

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

I link mothers to children for the analysis by mother's 1995 residence 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 household 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. This corrects for the fact that the matching process mistakenly assigns grandmothers as mothers in some cases, giving an unrealistic 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.

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, which is made up of 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 (2004): "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?," National Bureau of Economic Research Working Paper No. 10251.

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.

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," Working paper, UCLA.

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

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.

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.

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.

Page, S. E., F. Siegert, J. O. Rieley, H.-D. V. Boehm, A. Jaya, and S. Limin (2002): "The Amount of Carbon Released from Peat and Forest Fires in Indonesia in 1997," *Nature*, 420, 61–65.

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.

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.

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.

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.

Yount, K. M. (2001): "Excess Mortality of Girls in the Middle East in the 1970s and 1980s: Patterns, Correlates and Gaps in Research," *Population Studies*, 55(3), 291–308.

Figure 1: Map of Indonesia

Figure 2: Rainfall at Palembang Airport meteorological station, South Sumatra, 1990-97

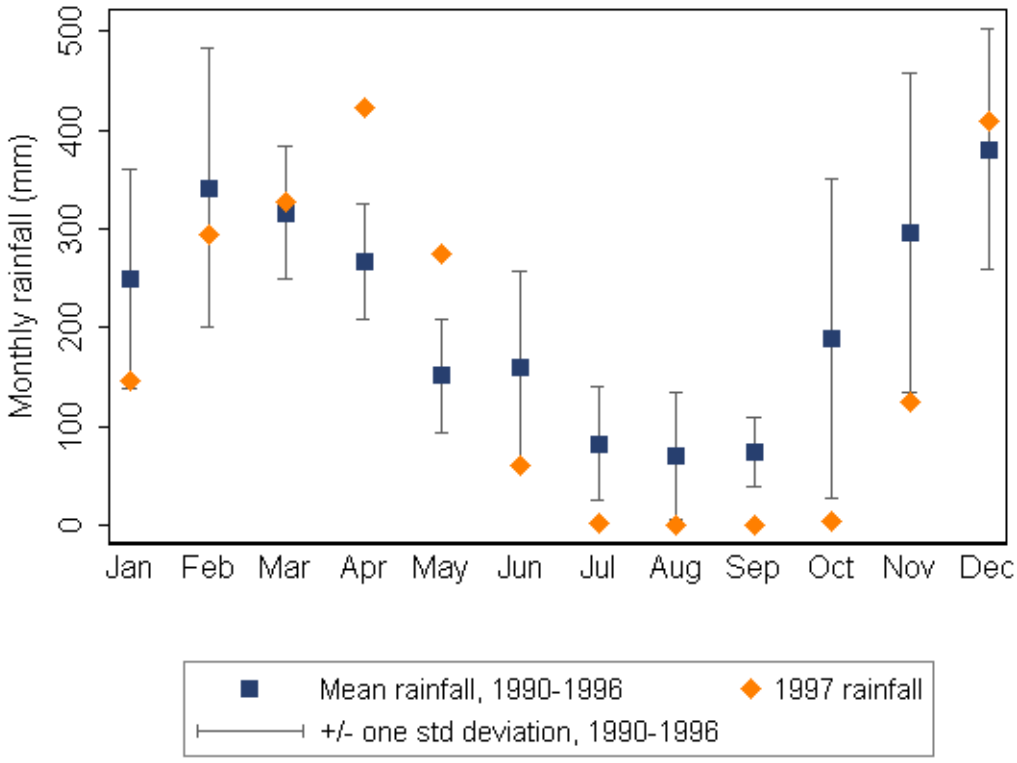
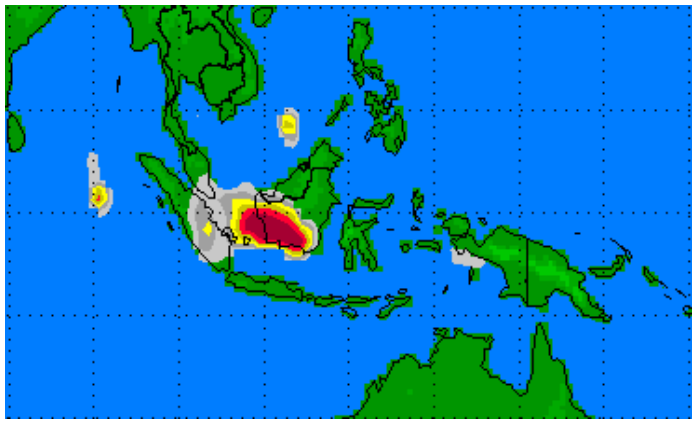
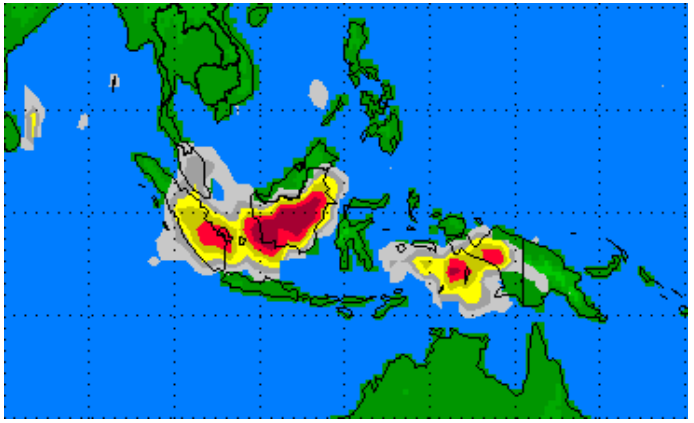


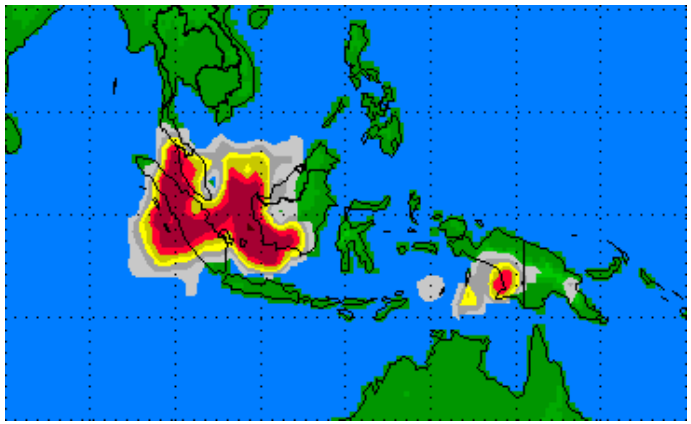
Figure 3: Satellite images of smoke over Indonesia



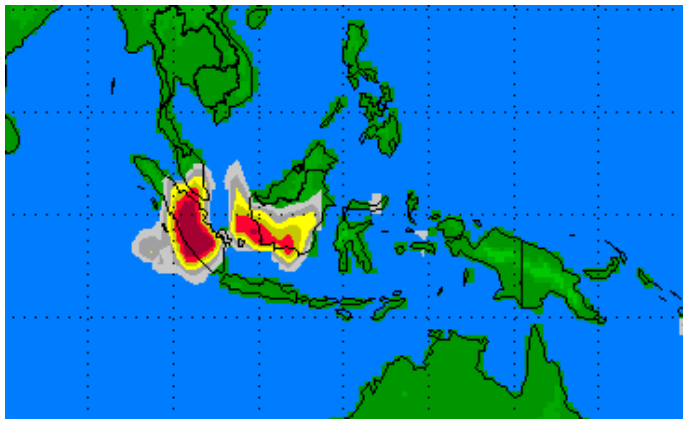
September 5, 1997



September 15, 1997

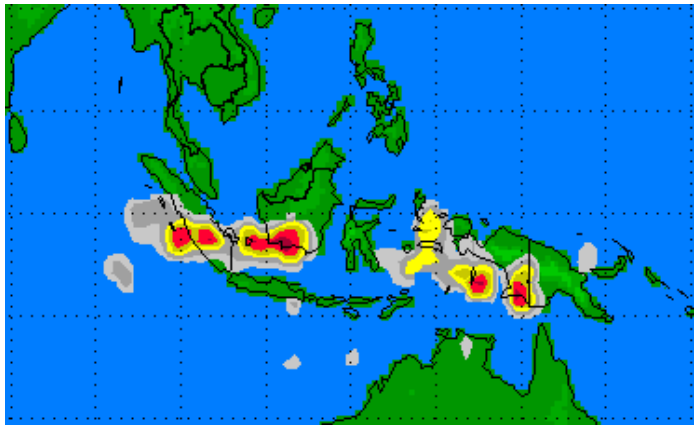


September 25, 1997

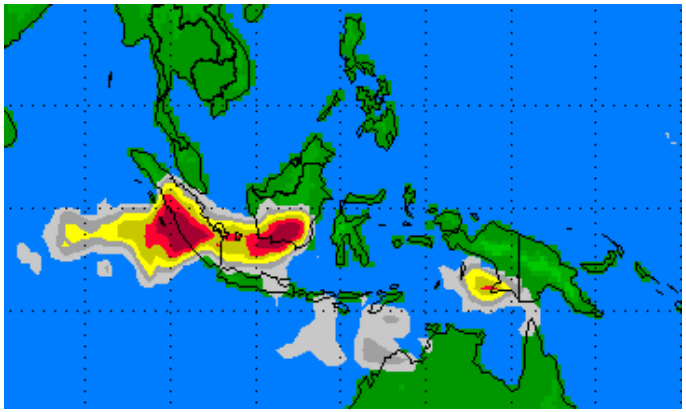


October 5, 1997

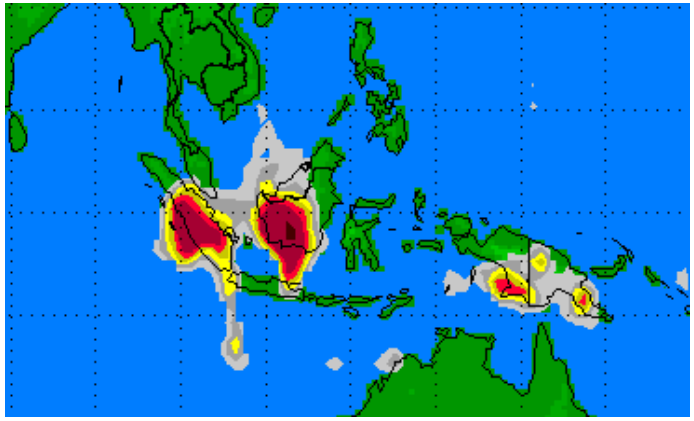
Figure 3 (continued): Satellite images of smoke over Indonesia



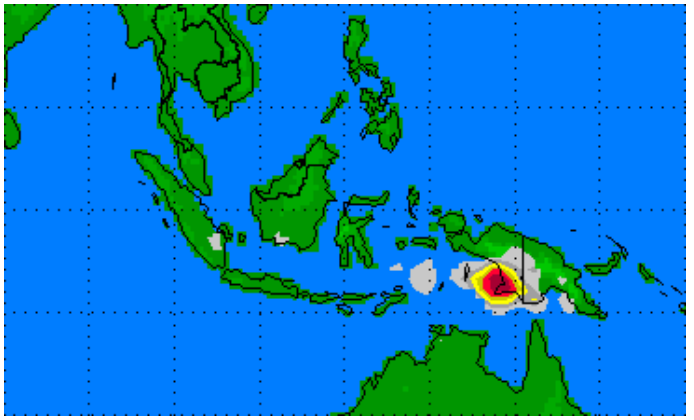
October 15, 1997



October 25, 1997



November 4, 1997



November 14, 1997

Figure 4: Timing and location of the pollution

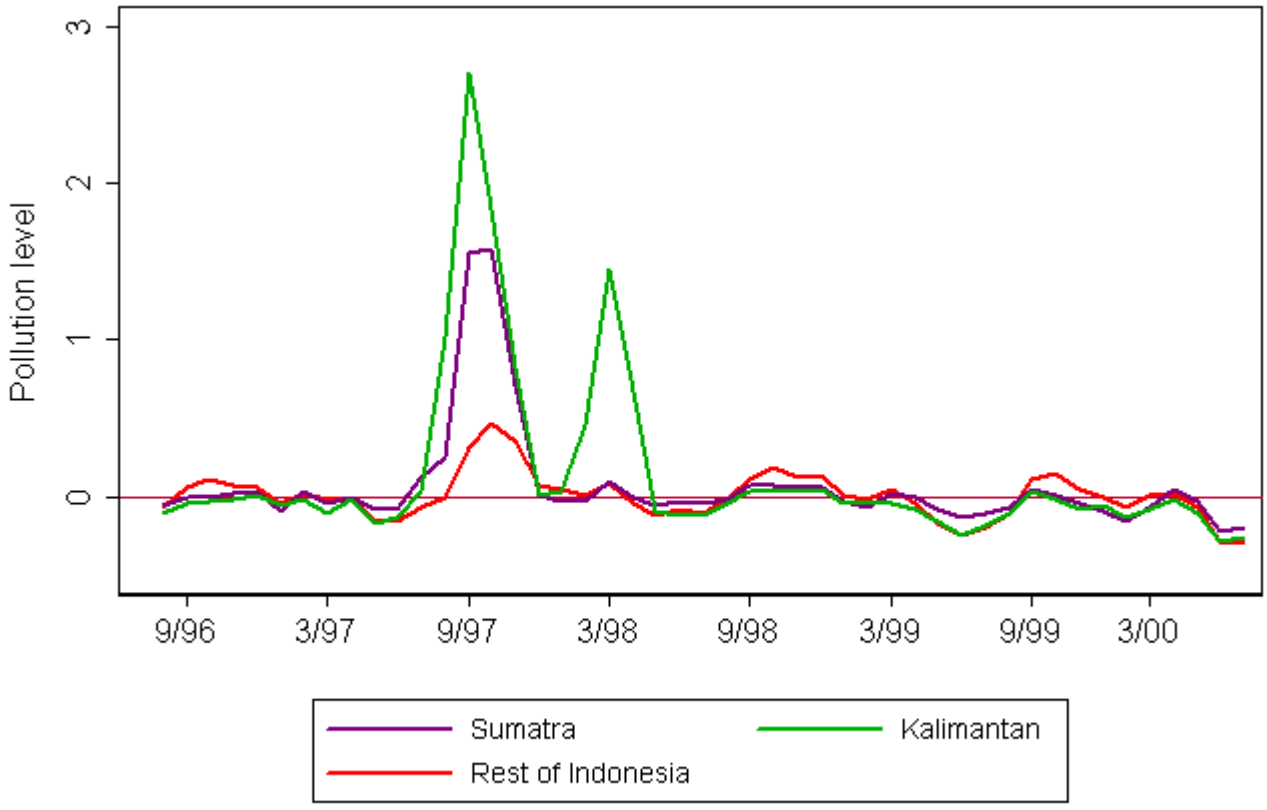
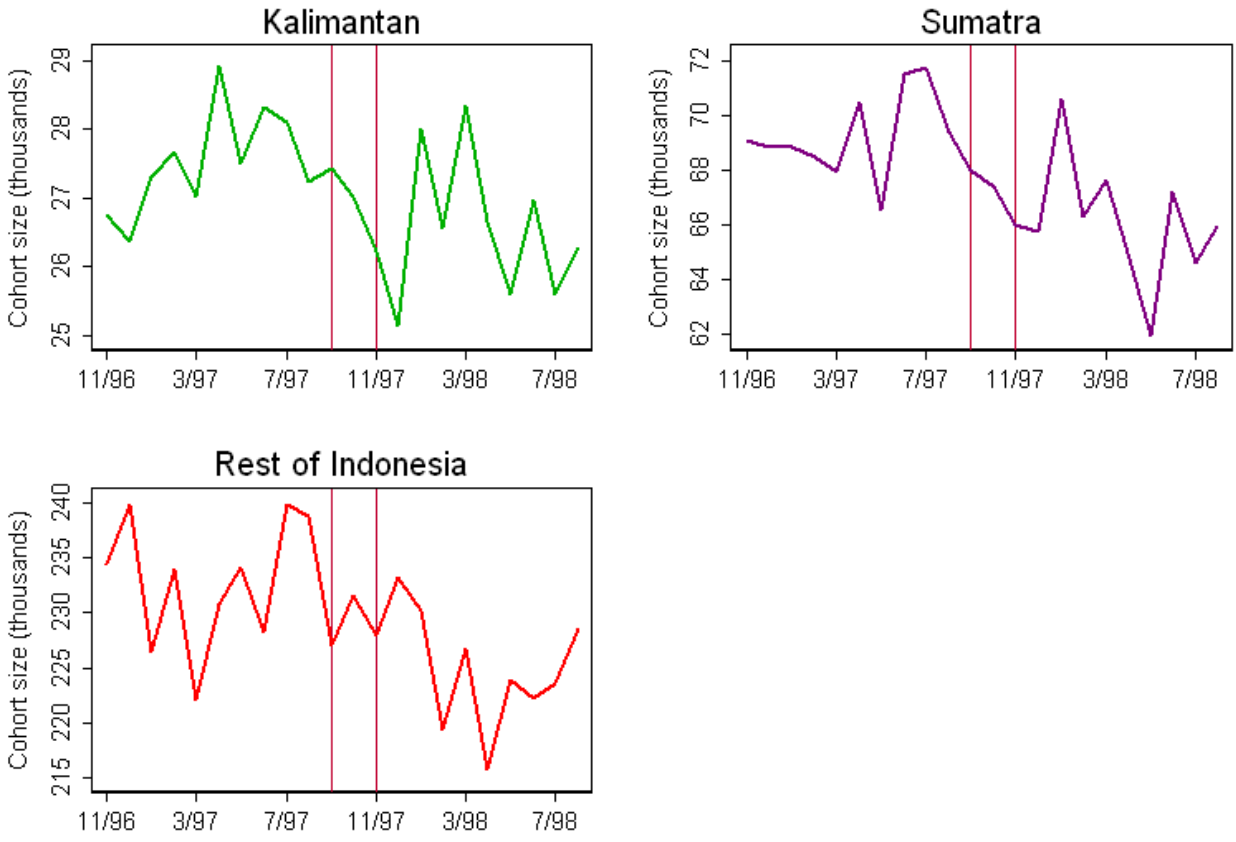
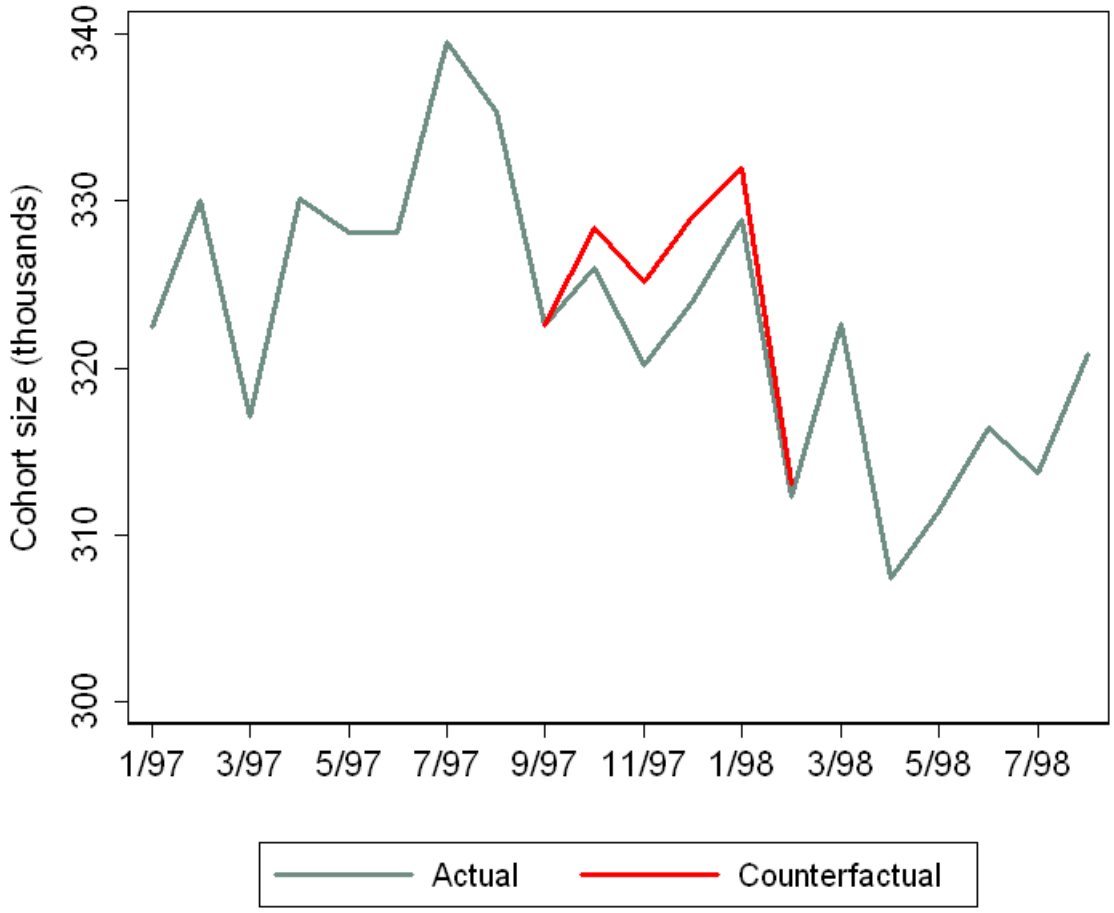


Figure 5: Fluctuations in cohort size by region



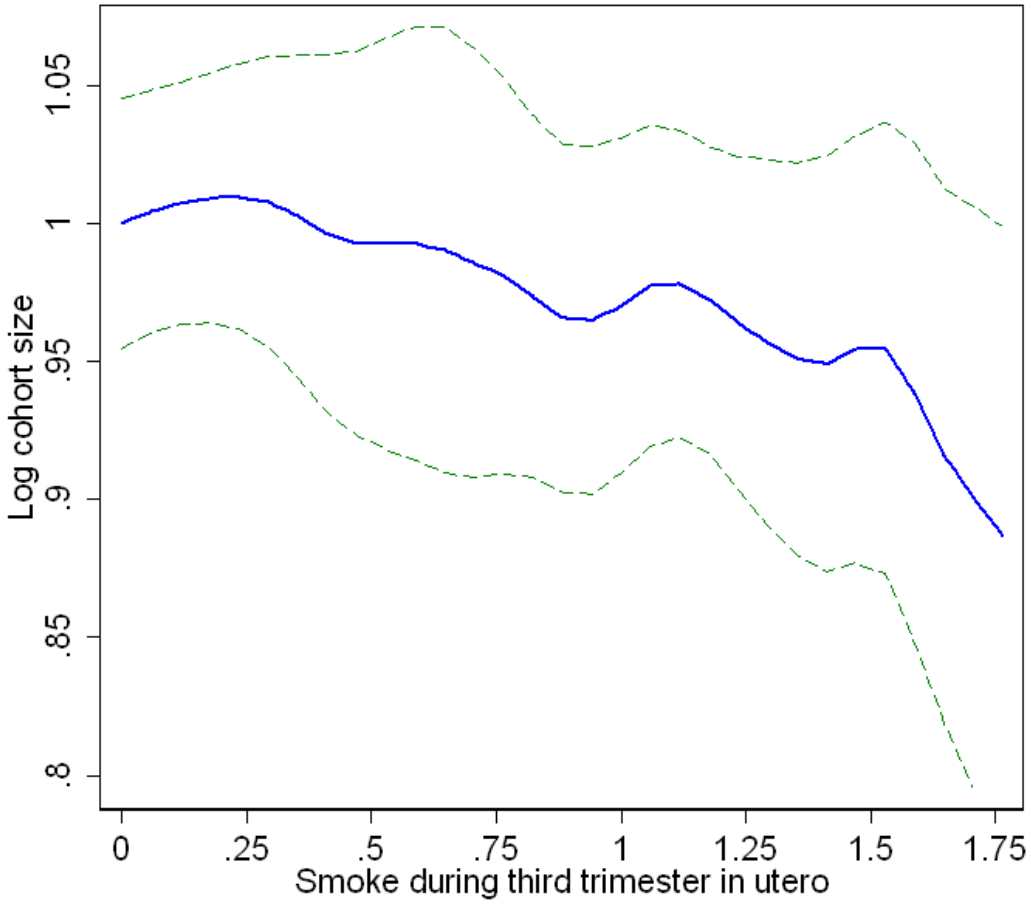
Cohort size is the total number of children enumerated in the Census, seasonally adjusted using the two years before the sample period (that is, mean differences in log cohort size by calendar month have been removed). The vertical bars demarcate September–November 1997, the period of the fires.

Figure 6: Missing children attributable to the fires



The actual cohort size is the total number of children enumerated in the Indonesian Census, by month of birth. The counterfactual uses the estimated coefficient for *PrenatalSmoke* (Table 3, column 2) and calculates what the population would have been if during and immediately after the fires, *PrenatalSmoke* had taken on its values from 12 months earlier. Cohort size is seasonally adjusted (mean differences in log cohort size by calendar month have been removed).

Figure 7: Kernel regression of log cohort size on third-trimester pollution



The solid line is the relationship between log cohort size and pollution ($PrenatalSmoke$). The dashed lines bound 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. The estimate is not weighted by population.

Figure 8: Timing of the fires and the financial crisis

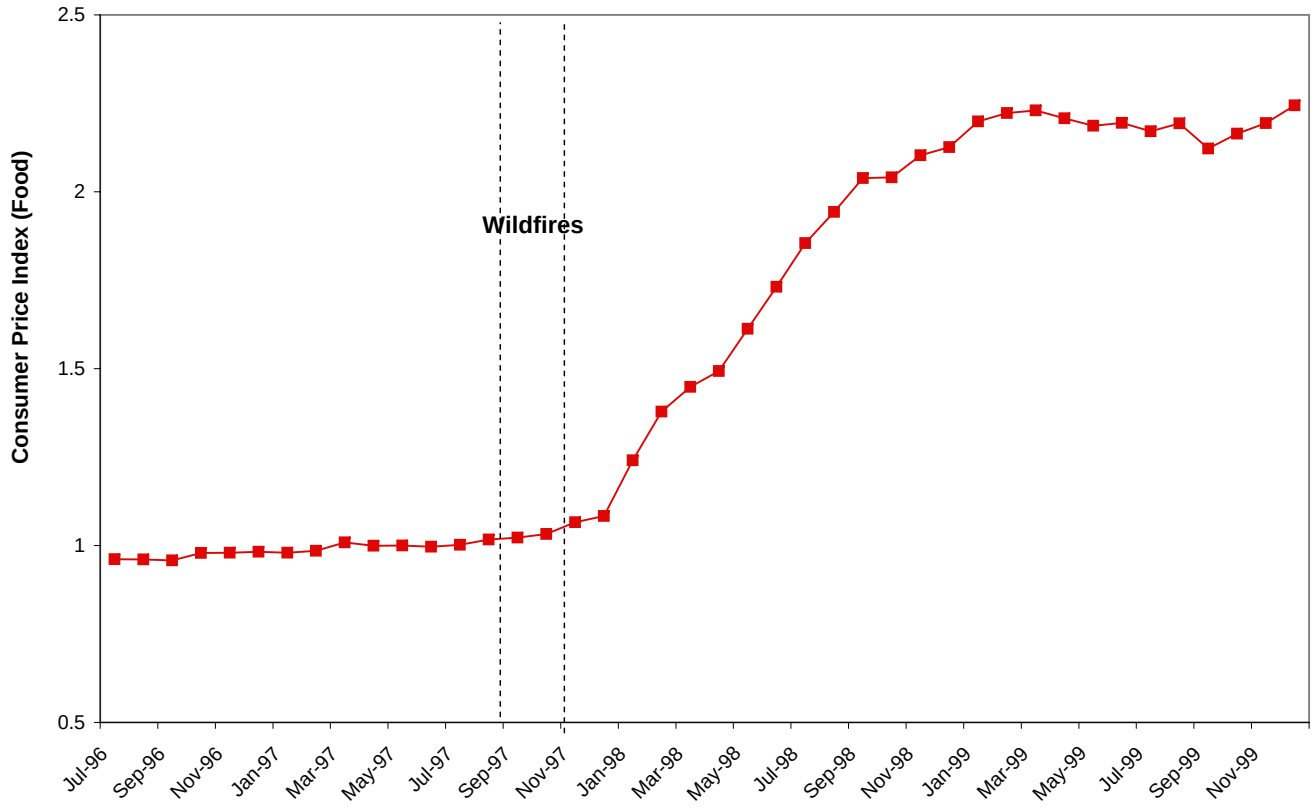


Table 1

Descriptive Statistics

	Mean	Std. Dev.
Cohort size variables		
Cohort size (for subdistrict-month)	95.6	89.7
Ln(cohort size)	4.8	.82
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 > .3)	.169	.262
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		
% urban	.57	.39
Ln(Predicted Births)	6.56	.91
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 (food)	1.131	.202
Wood as primary cooking fuel	.636	.413
Doctors per 1000 people	.161	.241
Midwives per 1000 people	.366	.261
Maternity clinics per 1000 people	.031	.050

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. Percent urban is 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. The data appendix describes how Ln(Predicted Births) is constructed using demographic information. 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 Dependent 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)	
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	.008	.001	.018	-.001
	(.006)		(.007)	(.030)	(.009)	(.014)	(.031)
Prenatal Smoke ($Smoke_{t-1,2,3}$)	-.035 ***	-.032 ***	-.032 **	-.089 **			
	(.012)	(.011)	(.013)	(.041)			
Postnatal Smoke ($Smoke_{t+1,2,3}$)	-.014		-.016 *	.001			
	(.009)		(.010)	(.034)			
$Smoke_{t-1}$					-.010	-.028 *	.006
					(.009)	(.016)	(.032)
$Smoke_{t-2}$					-.023 ***	-.006	-.036
					(.008)	(.013)	(.029)
$Smoke_{t-3}$					-.003	-.005	-.055 *
					(.013)	(.015)	(.030)
$Smoke_{t+1}$					-.010	-.019	-.022
					(.009)	(.014)	(.030)
$Smoke_{t+2}$					-.005	-.003	-.020
					(.008)	(.014)	(.027)
$Smoke_{t+3}$					.001	-.001	.002
					(.009)	(.012)	(.026)
Observations	67454	67454	67454	67454	67454	67454	67454
Subdistrict and month FEs?	Y	Y	Y	Y	Y	Y	Y

Notes: Each observation represents 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 birth versus mother's 1995 residence			Within district-month
	Residence	Birthplace	Mother's 1995 residence	
	(1)	(2)	(3)	(4)
Smoke	-.002	.002	.002	-.007
	(.006)	(.006)	(.006)	(.014)
Prenatal Smoke	-.035 ***	-.037 ***	-.038 ***	-.013
	(.012)	(.012)	(.012)	(.020)
Postnatal Smoke	-.013	-.015	-.016	-.0004
	(.010)	(.010)	(.010)	(.022)
Observations	5829	5829	5829	67454
Fixed effects	month, district	month, district	month, district	district*month, subdistrict

Notes: In columns 1 to 3, each observation is a district-month. In column 4, 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 5

Alternative Hypotheses

Dependent variable: Log cohort size

	Control for predicted fertility	Excluding September 1997	Control for financial crisis (no control, at birth and after birth)			Effects for areas with and without fires
	(1)	(2)	(3)	(4)	(5)	(6)
Smoke	.001	.0001	.002	.002	.002	.001
	(.006)	(.009)	(.006)	(.006)	(.006)	(.028)
Prenatal Smoke	-.035 ***	-.035 ***	-.032 ***	-.032 ***	-.033 ***	-.022
	(.012)	(.013)	(.011)	(.011)	(.011)	(.037)
Postnatal Smoke	-.014	-.013	-.012	-.012	-.013	.001
	(.009)	(.010)	(.009)	(.009)	(.009)	(.029)
Ln(Predicted Births)	.350					
	(.278)					
Smoke * Kalimantan						.005
						(.026)
Prenatal Smoke * Kalimantan						-.023
						(.036)
Postnatal Smoke * Kalimantan						-.022
						(.028)
Smoke * Sumatra						-.008
						(.026)
Prenatal Smoke * Sumatra						.002
						(.035)
Postnatal Smoke * Sumatra						-.003
						(.027)
Financial Crisis				-.049		
				(.038)		
Fin. Crisis in quarter after birth					-.023	
					(.029)	
Observations	67454	63703	63158	63158	63158	67454
Subdistrict and month FEs?		Y	Y	Y	Y	Y

Notes: Each observation represents 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.

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 represents 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)	(6)
Prenatal Smoke	-.121 ***	.015	-.115 ***	-.113 ***	-.130 ***	-.007
	(.028)	(.032)	(.027)	(.028)	(.028)	(.025)
Prenatal Smoke * Urbanization	-.013					
	(.013)					
Prenatal Smoke * Wood Fuel Use		-.155 ***				-.120 ***
		(.036)				(.026)
Prenatal Smoke * Matern. Clinic			.030 ***			.011 **
			(.009)			(.005)
Prenatal Smoke * Doctors				.048 ***		.016
				(.015)		(.013)
Prenatal Smoke * Midwives					-.006	
					(.009)	
Prenatal Smoke * High Consum	.071 ***	.048 *	.058 **	.052 **	.073 ***	.044 *
	(.027)	(.025)	(.025)	(.025)	(.028)	(.025)
Observations	63158	63158	63158	63158	63158	63158
Subdistrict and month FEs?	Y	Y	Y	Y	Y	Y

Notes: Each observation represents 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 if 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 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

Population and Mortality Statistics for Indonesia

	Population		Annual Pop Growth	Pop Density (km^{-2})	Infant Mortality Rate (per 1000 live births)				Under-5 Mortality Rate (per 1000 live births)			
	1990	2000	1990-2000	2000	1990	1994	1997	1998	1990	1994	1997	1999
Indonesia	179,378,946	206,264,595	1.49	109	71	66	52	49	99	93	71	60
Provinces												
Aceh	3,416,156	3,930,905	1.46	76	58	58	46	41	78	79	59	48
North Sumatra	10,256,027	11,649,655	1.32	158	61	61	45	43	82	97	72	52
West Sumatra	4,000,207	4,248,931	0.63	99	74	68	66	50	103	98	95	62
Riau	3,303,976	4,957,627	4.35	52	65	72	60	40	89	94	82	48
Jambi	2,020,568	2,413,846	1.84	45	74	60	68	45	102	88	82	55
Sumatera Selatan	6,313,074	6,899,675	2.39	74	71	60	53	51	98	92	70	62
Bengkulu	1,179,122	1,567,432	2.97	79	69	74	72	51	96	124	115	63
Lampung	6,017,573	6,741,439	1.17	191	69	38	48	49	96	58	64	60
Bangka Belitung	n/a	900,197	0.97	56	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
DKI Jakarta	8,259,266	8,389,443	0.17	12635	40	30	26	26	55	50	42	29
West Java	35,384,352	35,729,537	2.03	1033	90	89	61	56	129	120	77	69
Central Java	28,520,643	31,228,940	0.94	959	65	51	45	48	89	75	60	45
Yogyakarta	2,913,054	3,122,268	0.72	980	42	30	23	27	53	35	30	30
East Java	32,503,991	34,783,640	0.7	726	64	62	36	50	87	79	53	63
Banten	n/a	8,098,780	3.21	936	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Bali	2,777,811	3,151,162	1.31	559	51	58	40	33	67	63	44	38
W. Nusa Tenggara	3,369,649	4,009,261	1.82	199	145	110	111	85	216	160	150	114
E. Nusa Tenggara	3,268,644	3,952,279	1.64	83	77	71	60	59	108	108	90	75
W. Kalimantan	3,229,153	4,034,198	2.29	27	81	97	70	56	114	135	88	71
Central Kalimantan	1,396,486	1,857,000	2.99	12	58	16	55	33	77	38	69	38
S. Kalimantan	2,597,572	2,985,240	1.45	69	91	83	71	67	130	111	87	86
East Kalimantan	1,876,663	2,455,120	2.81	11	58	61	51	34	78	76	66	39
North Sulawesi	2,478,119	2,012,098	1.33	132	63	66	48	39	86	83	61	46
Central Sulawesi	1,711,327	2,218,435	2.57	35	92	87	95	63	132	127	121	80
South Sulawesi	6,981,646	8,059,627	1.49	129	70	64	63	38	97	86	79	45
SE Sulawesi	1,349,619	1,821,284	3.15	48	77	79	78	53	108	105	94	66
Gorontalo	n/a	835,044	1.59	68	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Maluku	1,857,790	1,205,539	0.08	26	76	68	30	48	107	91	48	50
North Maluku	n/a	785,059	0.48	25	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Irian Jaya	1,648,708	2,220,934	3.22	6	80	61	65	55	113	88	92	69

Source: Badan Pusat Statistik (Statistics Indonesia); based on Census and Demographic Health Surveys

Table A3
Different Sample Periods

Dependent variable: Log cohort size

	Shorter periods			Balanced calendar months		Longer period
	3/97 - 5/98	12/96 - 2/98	12/96 - 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 represents 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.