The Scourge of Asian Flu

In utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children

Elaine Kelly

A B S T R A C T

This paper examines the impact of in utero exposure to the Asian influenza pandemic of 1957 upon childhood development. Outcome data are provided by the National Child Development Study (NCDS), a panel study where all members were potentially exposed in the womb. Epidemic effects are identified using geographic variation in a surrogate measure of the epidemic. Results point to multiple channels linking fetal health shocks to childhood outcomes: physical development is impeded, but only when mothers had certain health characteristics; by contrast, the negative effects on cognitive development appear general across the cohort.

I. Introduction

The foundations for life-long health and human capital formation are laid in the womb. As the fetus develops, it is subject to a range of environmental influences, only some of which the mother can control. At the extreme, intrauterine exposures to certain maternal diseases, such as rubella, or drugs, such as alcohol, can lead to substantial physical and mental impairments at birth. More commonly,

Elaine Kelly is a research economist at the institute for Fiscal Studies. She recently completed her PhD at University College London. *She is indebted to Imran Rasul for his extensive advice and support. She is also grateful to James Banks, Steve Machin, Alissa Goodman, Wendy Carlin, Bob Michael, Doug Almond, Jim Smith, Ken Chay, Liz Washbrook, the participants of the "Persistence in Health" conference at Bristol University, and an anonymous referee, for their ideas, suggestions and comments. She thanks the ESRC and the IFS for financial support, and the University of London and U.K. Data Archive for access to the NCDS data. Local authority codes were obtained under UKDA special license (usage 29601). Lastly, she thanks the London School for Hygiene and Tropical Medicine for access to the Registrar Generals' Returns. She claims responsibility for all errors. The local authority level data used in this article, including pneumonia notifications and local area characteristics, can be obtained beginning May 2012 through April 2015 from Elaine Kelly, Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE, UK, e.kelly@ifs.org.uk. NCDS codes are available under license from the U.K. Data Archive. For information about how to apply for the license, please contact the author.*
[Submitted March 2010; accepted September 2010]
ISSN 022-166X E-ISSN 1548-8004

episodes or events may have impacts that are mild, remain unnoticed, or only emerge in later life. The original fetal origins (FO) hypothesis proposed a link between fetal nutrition and certain diseases in adulthood (Barker 1990). Current versions include a broader range of both potential hazards and relevant outcomes (Jaddoe and Witteman 2006).

The FO hypothesis is hard to test, as the substantial scope for omitted variables and nonrandom selection almost always preclude the use of nonexperimental data. The economics literature has thus used natural experiments to generate sharp, random shocks to fetal health conditions, either across space or cohorts. Examples include disease outbreaks (Almond 2006; Almond and Mazumder 2005), radioactive emissions (Almond, Edlund, and Palme 2008; Otake and Schull 1998), famine (Almond et al. 2007; Banerjee et al. 2007), fasting (Almond and Mazumder 2008), and policy-induced changes in alcohol consumption (Nilsson 2008). This literature has established that fetal health shocks can have significant effects upon physical and cognitive development. Yet, data are typically restricted to limited number of adult outcomes, recorded in repeated cross-sections. The full implications of the FO hypothesis for the broader economics of health and human capital formation literatures, and for public health policy, are thus hard to assess.

This paper examines the physical and cognitive development of a cohort of British children, in utero during the 1957 Asian influenza pandemic. Effects of exposure on outcomes at birth, and ages seven and 11 are identified using variation in the intensity of the influenza outbreak, across the 172 local authorities (LAs) of birth. The data provide a unique concurrence of a fetal health shock and a detailed panel of individuals, each of whom were potentially exposed. The objectives of the paper are two-fold: first, to establish whether Asian influenza did have significant effects upon childhood physical and cognitive development, a stage rarely observable to researchers; and, second, to use the panel structure of the data to investigate possible mechanisms behind the estimated relationships.

The National Child Development Study (NCDS) is an ongoing cohort study that follows all 17,400 individuals born in Great Britain between the third and ninth of March 1958. The first wave was conducted at birth, the second at age seven, and the third at age 11. Later outcomes are available, but are outside the scope of this paper. Waves 1–3 include detailed information on the development of the child, and the socioeconomic background and health of parents. The principal outcomes used are birth weight at birth, and cognitive test scores and height in cm at ages seven and 11. Asian flu struck Great Britain between September and November 1957, when the majority of NCDS cohort were in their second trimester. The epidemic is measured using a surrogate of the LA influenza infection rate, official pneumonia notifications per 100,000 population (Hunter and Young 1971).

Most existing evidence on the impacts of in utero exposure to pandemic influenza comes from the devastating Spanish flu of 1918–19. Perhaps the most comprehensive study is provided by Almond (2006), who exploits information on quarter of birth and outcomes in adulthood, as documented by the 1960, 1970, and 1980 U. S. censuses. He finds that cohorts exposed during the first or second trimester of pregnancy were significantly less likely to complete high school, had consistently lower earnings, received higher welfare payments, and had higher rates of incarceration, than predicted from trend. These findings are consistent with medical evidence on

the vulnerable periods for cerebral development, and with the remainder of the empirical FO literature (see, for example, Otake and Schull 1998; Almond et al. 2008; Guerri, Bazinet, and Riley 2009).[1] Results obtained using spatial variation in the virulence of the epidemic are poorly determined. However, Almond only has data for the nine U.S. census divisions, generating very little variation. By contrast, the NCDS has 172 geographic identifiers, covering a far smaller geographic area.

The contributions of this paper are two-fold. The first, but more minor contribution, is to estimate the impact of pandemic influenza in the early years, a period rarely observed in administrative data. The second, and more substantive contribution, is to provide some information on the mechanisms that might link fetal health shocks to subsequent outcomes. Two issues are addressed in detail: the extent to which the impacts of the epidemic on childhood outcomes are captured by birth weight; and, the role of maternal health in moderating the effects of influenza. There is a large literature that estimates the effects of birth weight upon subsequent physical and cognitive development (for example, Black, Devereux, and Salvanes 2007; Currie and Moretti 2007; Currie and Hyson 1999). However, a surfeit of potential omitted variables makes the mechanisms very hard to identify. Some studies have used birth weight as a proxy for the quality of the intrauterine environment and parental investment prior to birth (for example, Conley and Bennett 2000; Currie and Moretti 2007). The role of maternal health is particularly salient for the economic literature, as mother's health represents one potential channel for the intergenerational transmission of disadvantage, and advice for pregnant women can have important long-run public health implications. The main findings are as follows.

First, Asian influenza had a negative and statistically significant effect on birth weight, but only for the offspring of mothers who smoked before pregnancy or were short (154 cm or less). This pattern of heterogeneity is consistent with medical evidence on the possible relationships between influenza and fetal nutrition.

Second, statistically significant effects on test scores are present at the mean, with a one standard increase in epidemic intensity reducing scores by 0.07 standard deviations at seven, and 0.06 at age 11. These effects do not vary with cohort member characteristics, including the maternal smoking and height. Negative effects of influenza upon child height at ages seven and 11 are only present for smokers, replicating the results for birth weight. These patterns are consistent with two separate transmission mechanisms: (i) reduced physical development, where a mother's symptoms are plausibly severe, or when her nutritional stores are unable compensate for nutritional disruption; (ii) impaired cognitive development, irrespective of maternal characteristics, possibly as a consequence of the inflammation and hyperthermia that typically accompany influenza.

Third, the impacts of the epidemic on childhood height and test scores do not operate through, and are not captured by, birth weight. The association between birth weight and later physical and cognitive outcomes is well documented by the medical and social science literatures, but remains poorly understood (for example, Currie and Hyson 1999; Black, Devereux, and Salvanes 2007; Almond, Chay, and Lee

1. See Almond and Mazumder (2005) for results on health outcomes and Lin (2008), Nelson (2003), and Erikson, Sundet, and Tambs (2009), for international evidence.

2005; Godfrey and Barker 2001). Our results suggest that birth weight should not be viewed as a catchall measure of influences on prenatal development.

Fourth, maternal smoking plays a critical role in shaping whether influenza affected physical development, from birth onward. The effect of the epidemic upon birth weight and child height is negative only where mothers smoked prior to pregnancy. Our results thus reinforce the existing public health message that pregnant women should not smoke. The structure of the paper is as follows. Section II details the Asian influenza pandemic, and examines the possible links between maternal influenza and fetal development. Section III describes the data. Section IV presents the results for birth weight. Section V presents the results for child outcomes at ages seven and 11. Section VI details a series of robustness tests. Section VII discusses the implications of our results.

II. Background

A. The Epidemic

The Asian influenza pandemic of 1957–58 was the second of three twentieth century influenza pandemics. It was far milder than its predecessor, the catastrophic Spanish influenza of 1918–19, but claimed more lives than the Hong Kong influenza of 1968–69.[2] Each of these pandemics occurred when a new form of the Influenza A virus was introduced into the human population from a nonhuman—and in general, avian—host.

In nonpandemic years, influenza infection rates fluctuate with the seasons, claiming several hundred thousand lives worldwide each year. Pandemic influenza spreads in waves, without apparent regard for climate or season, and the death toll usually runs into the millions. While seasonal flu takes the heaviest toll on those younger than two and older than 65, victims of pandemic flu include high numbers of older children and prime-age adults.[3]

The risk of infection, the severity of symptoms, and death rates from influenza are all elevated during pregnancy, due to physiological changes in a woman's cardiovascular, respiratory, and immune systems. Those with underlying medical conditions or weakened respiratory systems are at a particularly high risk. Influenza rarely crosses the placenta to infect the fetus. Any negative effects on fetal development therefore operate through secondary responses, such as inflammation or an interruption in maternal nutrition, although the precise mechanisms remain poorly understood (Goodnight and Soper 2005; Arcavi and Benowitz 2004; Rasmussen, Jamieson, and Bresee 2008).

Asian influenza hit Great Britain between June 1957 and April 1958, but cases were concentrated between September and November 1957, henceforth known as

2. U.S. government estimates suggest that Spanish flu killed 40 million worldwide, compared with 1–2 million for Asian flu, and 0.7 million Hong Kong flu (DHHS 2008).

3. In nonpandemic years, those younger than 65 account for between 10 and 20 percent of all influenza deaths, as compared 36 percent during Asian Flu, 48 percent during Hong Flu and 99 percent during Spanish Flu (Simonsen et al. 1998).

the "epidemic period." In 1957 alone, the epidemic was responsible for 30,000 excess deaths in England and Wales, of which 6,716 were from influenza.[4] According to official estimates, at least six million people in England and Wales (or 13 percent of the population) suffered some level of incapacitation during the main epidemic period. Cases were concentrated among those aged between four and 39, with the highest incidence among those of school age. For women of child-bearing age, the infection rate was approximately 30 percent, but there no separate figures for pregnant women (Woodall, Rowson, and McDonald 1958; Ministry of Health 1960).

Influenza was not a notifiable disease in 1957, so the total incidence is only available by proxy. We adopt the approach of Hunter and Young (1971), and use official pneumonia notifications made to the Registrar Generals of England and Wales, and of Scotland. Pneumonia is closely clinically related to influenza, an attack of influenza may result in pneumonia, and quarterly trends in pneumonia notifications closely replicate influenza deaths. The pattern of weekly pneumonia notifications shows that the epidemic followed an approximate normal distribution. Notifications departed from their long-run trend at the beginning of September, and returned to normal at the end of November; the peak occurred in the week ending October 19, 1957 (Hunter and Young 1971).

The effect of the epidemic is identified using the spatial variation in pneumonia notifications per hundred thousand population, across the 172 local authorities of birth. The identifying assumption is that the departure of notifications from their long-run trend is exogenous. Infection rates were higher in Northern England and Scotland, but there were significant variations within regions. The epidemic peaked in the week ending October 19, when 90 percent of the NCDS children were between 17 and 23 weeks in gestation.

B. In Utero Exposure to Influenza and Child Outcomes—Possible Channels

1. Maternal Health Shocks, Influenza and Birth Weight

Birth weight is used by medical, scientific, and social science literatures as a proxy for the quality of conditions for growth and development in the womb (Conley and Bennett 2000; Currie and Moretti 2007). It is an object of interest only in so far as it may linked, by cause or correlation, to subsequent health or socioeconomic outcomes (Jaddoe and Witteman 2006, page 93). There is nothing inherently good about heavier births.[5]

Birth weight is determined by two factors: gestation and intrauterine growth. A baby born before 37 weeks is defined as premature, and is at risk of health complications caused by incomplete development. The cause of premature birth is often unknown: gestation is unresponsive to policy interventions, such as improving access to healthcare or providing nutritional supplements, and is not that sensitive to maternal behavior (Kramer 1987; Goldenberg and Rouse 1998; Godfrey and Barker

4. By way of comparison, excess mortality during the Spanish Flu stood at 200,000, with 150,000 deaths attributed to influenza (Ministry of Health 1960).

5. Indeed, very high birth weight has been found to increase the risk of breast cancer and to impair cognitive development (Dos Santos Silva, De Stavola, and McCormack 2008; Cesur and Rashad 2008).

2001). Existing evidence suggests that any link between pandemic influenza and premature birth is attributable to complications, rather than the virus itself (Rasmussen, Jamieson, and Bresee 2008).

Impeded intrauterine growth (IUG), or fetal malnutrition, occurs when the birth weight of a child born at term falls below a certain threshold; this, by definition, is caused by insufficient nutrient uptake in the womb.[6] The fetal origins hypothesis postulates that fetal under-nutrition can lead to "reprogramming," or altered gene expression, which has permanent impacts upon an individual's physiology (Jaddoe and Witteman 2006). Consequences include a higher risk of strokes, cardiovascular disease, and diabetes in adult life. IUG has proved responsive to policy interventions designed to improve maternal nutrition during pregnancy and is impaired by smoking (Kramer 1987; Frisbie et al. 1997; Goldenberg and Rouse 1998; Godfrey and Barker 2001).

The nutrients a fetus receives from its mother come from two sources: her diet during pregnancy, and the stock of nutrients stored in her liver and the tissues of her body. Barker (2009) claims that it is latter, the mother's nutrient reserves, that are of primary importance for fetal growth; this, he suggests, can be proxied by her height. Nutrients are delivered to the baby through the placenta; the efficiency of these transfers can vary with placental size, which is in turn affected by maternal nutrition.

Maternal influenza could affect fetal nutrition flows in at least three ways: First, through suppressing appetite and reducing nutritional intake, the consequences of which may be more severe in pregnant women, due to biochemical changes in their bodies (Metzger et al. 1982). Secondly, by increasing excretion rates, and interfering with the absorption of fats, proteins, and other essential nutrients. Lastly, through inducing fever, which acts to accelerate energy consumption and nutrient loss (Tomkins et al. 1994). The rate of nutritional depletion depends on the characteristics of the mother; in particular, the diminution is more rapid where the preinfection nutrient intake was low (Scrimshaw 1977, page 1538; Tomkins et al. 1994; Edwards 2007; Scrimshaw 1977).[7]

2. Maternal Health Shocks and Brain Development

The human brain is more susceptible to teratogenic insults than most other embryonic and fetal structures; vulnerability is highest between eight and 25 weeks in gestation, when the brain is developing (Nyagu et al., 2002, p. 202). As a consequence, significant negative effects of fetal shocks, such as exposure to influenza, rubella, alcohol, and radiation, are only found during this vulnerable period (Almond 2006; Otake and Schull 1998; Almond et al. 2008; Lambert 2007; Guerri, Bazinet,

6. The most precise measure uses weight/meters^3; Barker (2009) specifies a threshold of 26 kg/m^3. Where a child's dimensions are unavailable, weight is adjusted for gestation, but exact thresholds vary from paper to paper (Martorell and Gonzalez-Cossio 1987).

7. Approximately one-third of influenza cases are asymptomatic, and the impact upon fetal development of these infections is unclear (Carrat et al. 2008). Our LA proxy is assumed to measure the rate of infections that do have the potential to affect fetal development.

and Riley 2009). More than 99 percent of the NCDS cohort members were between eight and 25 weeks in gestation when the epidemic peaked in their LA birth.

Brain development takes place over two critical periods. The first occurs between eight and 15 weeks, when the proliferation of neuronal (nerve cell) elements hit its peak and there is substantial migration of neurons to different parts of the developing brain. A fetal health shock during this period has the most severe and pronounced effects upon cognitive development (Otake and Schull 1998).

The second critical period takes place between 16 and 25 weeks: rapid neuron differentiation sees neurons developing different, and specific, biochemical and physiological properties; synaptogenesis, or the creation of new synapses hits its peak; the architecture of the brain begins to form; and, neuronal pruning takes place, eliminating more than 50 percent of neurons prior to birth (Otake and Schull 1998; Nyagu et al. 2002). Some evidence (for example, O'Callaghan et al. 1991) suggests that exposure to influenza at this stage interrupts the neuronal pruning process, increasing the child's susceptibility to schizophrenia in adulthood. Possible causes of this interruption are influenza-induced hyperthermia and inflammation (Rasmussen, Jamieson, and Bresee 2008).

The majority of NCDS intrauterine exposures to Asian influenza would have occurred during the second critical period of brain development. The epidemic peaked in the week ending the 19th October, when 95 percent of the NCDS cohort were between 16 and 25 weeks in gestation. Allowing a two-week infection window either side of the peak in each LA implies that a minimum of 80 percent of exposures 16 and 25 weeks. The remainder of exposures would have occurred almost exclusively between eight and 15 weeks.

III. Data

A. The National Child Development Study

The National Child Development Study (NCDS) is an ongoing longitudinal cohort study, which follows all those born in England, Scotland and Wales between the third and ninth March 1958, a total of around 17,400 births. Public use data files are augmented by LA identifiers for the first four waves under special license (University of London. IoE 2008b). Great Britain is divided into 172 local authorities; populations range from 18,400 (Shetland) to 5.5 million (London), with a median of 134,000. Observations are dropped entirely if birth weight is missing, reducing the total sample to 16,765.[8]

The NCDS began with the Perinatal Mortality Survey of 1958 (PMS), which was designed to examine the social and obstetric factors associated with stillbirth and death in early infancy. There have been eight subsequent followups, the most recent of which was completed in 2008. We use data from the first three waves, collected

8. Those with missing birth weight are present in later waves, and are spread across 124 of the 172 local authorities. In the 1958 survey, this group had a higher nonresponse rate to other survey questions, but, where background characteristics are recorded, they are not statistically different from the rest of the sample. In the 1965 and 1969 surveys, responses and nonresponses do not differ from the rest of the sample.

in 1958, 1965 and 1969, at birth, and ages seven and 11 (University of London, IoE 2008a). Waves 1–3 contain information on both parents and children. Parental data include health status, behavior, anthropometric measures, and socioeconomic background. Cohort member data include health outcomes, anthropometric measurements, cognitive test scores, and teacher behavioral and aptitude assessments. Participation fell from 16,765 in 1958 to 14,358 in 1965 and 14,069 in 1969.[9]

The principal outcome used from the PMS is birth weight, in grams, which has a mean of 3295.4 and a standard deviation of 579.7. Gestation in days is defined for 89.5 percent of the sample, distributed with a mean of 280 and a standard deviation of 14. Intrauterine growth can be calculated by adjusting birth weight for gestation.[10]

A series of cognitive tests were administered in both 1965 and 1969. We use the Draw a Man Score, aged seven, and the General Test Score (nonverbal), taken at age 11. These tests are chosen for having a larger range than others at the same ages, and distributions which are close to normal. However, the tests are not of the same form; there is therefore no way to distinguish changes in performance over time from differences in the test.[11]

The Draw a Man test asks children to draw a man; grades are based on detail and accuracy. The General Test was designed to test general ability through verbal and nonverbal reasoning. Absolute test scores are thus normalized, to follow a standard normal distribution. Heights, in cm, are measured in the same years as the tests.

B. The Epidemic

In absence of mother specific infection data, the epidemic is measured at the finest level of geographical aggregation available in the NCDS: the LA of birth.[12]

The epidemic is measured by the ratio of total number of pneumonia notifications during the epidemic period (September–November 1957) to population in LA of birth.[13] This acts as a measure for the likelihood that the cohort child's mother was infected by influenza, or indirectly affected by the infection rate in the local area.

9. Attrition was not additive, as 15,265 participated in either 1965 or 1969.

10. The mean birth weight of those with missing gestation is significantly lower than the rest of the sample (mean 3242.5, standard deviation 605.8). Their parents are also, on average, less educated and of a lower social class, suggesting possible scope for missing gestation on the basis of unobservables.

11. The NCDS administered reading and math tests in all three waves, but the distribution of results differ across tests and rarely approximate the normal. Conditional on wave participation, 95 percent took the Draw a Man test and 92 percent took the General Test.

12. As part of the PMS, the mothers were asked whether they had certain illnesses when pregnant, including influenza. These data do not appear in the files held by the U.K. data archive or the Centre for Longitudinal Studies.

13. The assumptions made when using pneumonia notifications to proxy for influenza infectious rates are as follows: (i) Physicians diagnoses were correct; (ii) Reportage was complete; (iii) Errors would occur randomly if some diagnoses were erroneous or some reportage were incorrect, and thus relative differences between local authorities would not be invalidated; (iv) Acute pneumonia is clinically associated with influenza in an acceptable parameter relationship. Hunter and Young (1971) propose that the ratio of pneumonia notifications to influenza infections is one to 417; (v) Subclinical cases of influenza need not be taken into account (Hunter and Young 1971, page 642); (vi) Any difference between LA of birth and LA of exposure is random with respect to the epidemic. If assumptions one and two fail to hold, the epidemic's effects will be attenuated toward zero.

Pneumonia notifications are given at a LA level, by week and by quarter, in the Registrar General's Returns for England and Wales, and in the corresponding Returns for Scotland (Registrar General for Scotland 1957; Registrar General for England and Wales 1957). LA population estimates are provided by the same source, as of first June 1957. Deaths from pneumonia or influenza are not given by week or quarter until 1958.

Further figures are collected on the number of pneumonia notifications in the same period in 1955 and 1956. Totals from these two years are averaged and divided by the LA population in 1956. Preepidemic intensity captures the underlying level of nonpandemic pneumonia in each LA. Other LA control variables come from the 1956 Registrar Generals' Returns and the 1951 census.

The descriptive statistics of all epidemic measures and LA controls are described in Table 1. Pneumonia notifications during the epidemic period ($Epid_l$) were four times higher than those in the previous two years ($PreEpid_l$), at the mean, the median, and for the interquartile range.[14] The last column shows the correlation between excess cases in 1957 ($Epid_l$–$PreEpid_l$), and the epidemic measures and LA controls. Excess notifications are almost perfectly correlated with epidemic measure, $Epid_l$ (0.96); the correlation with $PreEpid_l$ is far lower (0.46), indicating substantial variation in the intensity of the epidemic, conditional on the underlying prevalence of pneumonia. Of the LA controls, the highest correlate is population density (0.61), followed by percent Crowded (0.41). The last line of the table gives summary statistics for pneumonia notifications in 1958. Differences between $PreEpid_l$ and $PostEpid_l$ are small, and not statistically significant. $Epid_l$ therefore represents a large deviation from the preepidemic baseline, which was specific to the epidemic period. Nevertheless, $PostEpid_l$ is omitted from all specifications, so as avoid any possible endogeneity.

IV. Birth Weight

A. Empirical Method

The effect of the epidemic on the birth weight of child i, in LA_l, is estimated using the following linear specification:

$$(1) \quad BW_{il} = \alpha + \beta Epid_l + \theta PreEpid_l + \gamma LA_l + \chi X_{il} + \varepsilon_{il}$$

where $Epid_l$ represents the number of pneumonia notifications per hundred thousand population in the child's LA of birth; $PreEpid_l$, the same measure in the two years previous; LA_l, LA characteristics; and, X_{il} child-level characteristics at birth. The error term, ε_{il}, is assumed to be conditionally uncorrelated with $Epid_l$, and is clustered at the LA level. Equation 1 only considers the effects of the epidemic at the mean; interactions between $Epid_l$ and maternal characteristics are subsequently added to

14. The median and interquartile ranges of $Epid_l$ are 29.1 and 29.3, as compared with 7.24 and 7.17 for $PreEpid_l$. Both $Epid_l$ and $PreEpid_l$ have large standard deviations (36.7 and 10.4). Glasgow is an outlier in both $Epid_l$ and $PreEpid_l$. Dundee is an outlier in the $Epid_l$ measure, but $PreEpid_l$ value falls within the normal range.

Table 1
Summary Statistics of Maternal Characteristics Used to Test for the Heterogeneity of Epidemic Effects across Cohort Members

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum	Correlation with ($Epid_l - PreEpid_l$)
Pneumonia notifications per hundred thousand population (September-November 1957) ($Epid_l$)	172	39.71	36.72	0.89	261.72	0.96
Average pneumonia notifications per hundred thousand population, September-November 1955 and 1956 ($PreEpid_l$)	172	9.64	10.95	0	106.69	0.46
Still births/total births (1956) ($StillbirthRate_l$)	172	0.02	0.01	0.01	0.09	−0.05
Percentage of households living with >1 person per room ($PctCrowded_l$)	172	0.2	0.1	0.08	0.51	0.41
Proportion of male working population unskilled ($PctUnskilled_l$)	172	0.13	0.04	0.07	0.3	0.17
Proportion of men leaving school aged 16 or older ($PctSchoolPost16_l$)	172	0.13	0.05	0.04	0.45	−0.17
Population in thousands/square km ($PopDensity_l$)	172	1.81	1.96	0.01	7.91	0.61
Proportion of population 65 plus ($Pct65plus_l$)	172	0.11	0.023	0.056	0.2	−0.33
Pneumonia notifications per hundred thousand population, September–November 1958 ($PostEpid_l$)	172	8.74	9.43	0	87.12	0.51

Notes: $Epid_l$, $PreEpid_l$, $PostEpid_l$, $StillbirthRate_l$ and $DRTB_l$ are calculated using the Registrar General's Returns for England and Wales, and for Scotland. $Epid_l$ is calculated by dividing the total number of pneumonia notifications in the local authority in September, October and November 1957 by the local authority population, as estimated in June 1957. $PreEpid_l$ takes an average of the number of pneumonia notifications in a local authority in the same periods in 1956 and 1955, and divides by the population in 1956. $PostEpid_l$ gives the corresponding notifications/population figure for 1958. $PctCrowded_l$, $PctUnskilled_l$, and $PctSchoolPost16_l$ are calculated using the 1951 census 10 percent sample tables. $Pct65plus_l$ is calculated using the original local authority census tables. $PopDensity_l$ is calculated using local authority geographic areas, available from *UKBORDERS*, and the 1956 population estimates from the Registrar's Returns.

allow for heterogeneity across cohort members. For ease of interpretation, we refer to the magnitude of coefficients in terms of implied effect size, $\hat{\beta}\left[\frac{sd(Epid_l)}{sd(BW_{il})}\right]$: the marginal effect of a one standard deviation increase in $Epid_l$ on BW_{il}, in standard deviations.

LA controls, LA_l and $PreEpid_l$, are vital to ensuring that $\hat{\beta}$ is not biased by omitted variables at the local authority level. The controls, described in Table 1, attempt to capture features of the local environment or population that are potentially correlated with child outcomes, and could increase the rate of infection or make symptoms more severe. $PreEpid_l$ controls for the underlying rate of pneumonia in the area, so that $\hat{\beta}$ estimates the effect of the excess notifications during the epidemic period.

The vector X_{il} contains parental and background controls for each child, including the social class and schooling of both parents, mother's age and its square, mother's height, tenure of accommodation, number of persons per room, and whether the mother smoked before pregnancy. These characteristics are assumed to be predetermined, but could plausibly influence the probability of infection. Child-level information will later be exploited to test for heterogeneous effects across parental characteristics.

The coefficients estimated in Equation 1 should represent a lower bound of the true effects of Asian flu. $Epid_l$ is measured at a LA level; the estimated coefficients are thus a weighted average of the children of mothers who were affected by the epidemic and those who were not. The fetal origins literature often approximates effects on treated individuals by multiplying estimates by the inverse of the population exposure rate (for example, Almond 2006; Banerjee et al. 2007). Official estimates suggest one-third of women of child-bearing age contracted Asian flu, but the infection rate among pregnant women could be higher. Moreover, any adjustment factor would not account for heterogeneity across mothers, or any spillovers upon those not infected. Estimates are therefore left unadjusted.

B. Results

Table 2 presents the estimated impacts of Asian influenza upon birth outcomes. The first three columns consider the effect of the epidemic upon mean birth weight, in grams. Column 1 presents the unconditional effect of $Epid_l$, which is negative and significant at the 1 percent level. Introducing LA controls in Column 2 causes the estimated $\hat{\beta}$ to double in magnitude, the opposite effect to that expected. However, when child-level controls are introduced in Column 3, the coefficient is cut back in half and is no longer statistically significant. The pattern is the same for both gestation and IUG.[15]

Although the epidemic measure has no effect upon the mean, the baseline specification could mask heterogeneity across cohort members. With $Epid_l$ held constant, such variation could arise through differences in the rates of infection of pregnant women, or variation in the severity of the symptoms. In general, it is not possible

15. There is no statistical difference between the mean birth weight of the NCDS cohort, and that of the British Cohort Study, born 12 years later, a result consistent with an absence of an effect of the epidemic on mean birth weight.

Table 2
Birth Weight, in Grams, and the Intensity of the Epidemic

	(1) No Controls	(2) LA controls	(3) LA + Child Controls	(4) Birth Weight Interactions
$Epid_l$	−0.273***	−0.472**	−0.195	0.159
	(0.0722)	(0.200)	(0.210)	(0.306)
$PreEpid_l$		0.0533	−0.380	−0.429
		(0.518)	(0.542)	(0.570)
Smoking—nonsmoker omitted				
$Epid_l$ × smoking <10/day				−0.287*
				(0.161)
$Epid_l$ × smoking 10+/day				−0.411**
				(0.167)
Maternal height (65–66 inches, 164–169 cm) omitted				
$Epid_l$ × <=60 (<=154 cm)				−0.553**
				(0.223)
$Epid_l$ × 61 (154–56 cm)				−0.282
				(0.241)
$Epid_l$ × 62 (156–59 cm)				0.178
				(0.233)
$Epid_l$ × 63 (159–61 cm)				−0.0929
				(0.250)
$Epid_l$ × 64 (161–64 cm)				−0.117
$Epid_l$ × >=67 (>169 cm)				−0.257
				(0.574)
LA controls	No	Yes	Yes	Yes
Child-level controls	No	No	Yes	Yes
P-values—joint significance tests for $Epid_l$				
No effect of $Epid_l$	0.0002	0.0193	0.3544	0.0039
No difference by maternal smoking				0.0391
No difference by maternal height				0.0044
Observations	16,765	16,765	16,765	16,765
R-squared	0.0004	0.004	0.082	0.083

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent level. The dependent variable is birth weight in grams. Observations are at the cohort member level. Robust standard errors are clustered at the local authority level. $Epid_l$ represents pneumonia notifications per 100,000 population in LA of birth September-November 1957; $PreEpid_l$ is an average of notifications per 100,000 population in the same periods in 1955 and 1956. Local authority controls come from the 1956 Registrar Generals' returns (rates of still birth and death from TB, and population density) and from the 1951 census (percentage of men leaving school aged 16+, percentage of working age men unskilled, the proportion aged 65 and older, and percentage of households living with >1 person per room). See Table 1 for descriptive statistics. Child-level controls include social class and schooling of both parents, mother's age and its square, mother's height, the tenure of accommodation and number of persons per room, and whether the mother smoked before pregnancy.

to distinguish between the two, and it is not the aim of this paper to do so. The intention is instead to assess whether Asian influenza had an effect on the birth weight of any groups of cohort members, and to evaluate whether any patterns are consistent with the evidence on fetal nutrition presented in Section IIB1.

In Columns 4, $Epid_l$ is interacted with two indicators for maternal health: her height and prepregnancy smoking behavior.[16] Both are predetermined, and are thus left unaltered by the epidemic. Maternal height can be used as at proxy for life-long nutrition, and thus the stock of resources from which the fetus can draw upon (Barker 2009). Smoking is an established cause of restricted IUG, and could increase both the risk of influenza infection and the severity of symptoms (NHS 2009).

The results in Column 4 indicate that Asian influenza did significantly reduce birth weight, but only when mothers were short (154 cm and shorter) or who smoked prior to pregnancy.[17] For short mothers the implied effect of the epidemic, $\hat{\beta}\left[\frac{sd(Epid_l)}{sd(BW_{il})}\right]$, is to reduce birth weight by 0.04 standard deviations, relative to children of mothers 164–69 cm tall. Implied effects when mothers smoked were -0.03 for heavy smokers and -0.02 for light smokers. Both sets of interaction coefficients are approximately monotone, with absolute magnitudes falling as amount smoked decreases and height increases. These are precisely the relationships predicted by the medical evidence presented in Section IIB1. Taken together, all epidemic measures are jointly significant at the 1 percent level. When birth weight is decomposed into gestation and IUG (birth weight controlling for gestation), results suggest that the variation by maternal height is statistically significant for gestation, while heterogeneity by smoking is statistically significant for IUG. However, the preferred specification is that Column 4, as gestation has missing values that are potentially nonrandom, and has no effect on the child outcomes used in Section V, conditional on birth weight.[18]

In contrast to maternal health, there is no significant interaction between the epidemic and socioeconomic variables, such as mother's education or father's social class. Equally the effect of the epidemic does not vary other LA characteristics, such as population density, or with the cohort member's gestation. Block bootstrapped quantile regressions indicate that there is no quantile of the birth weight distribution where $Epid_l$ is significant, at the 5 percent level. A series of robustness tests is presented in Section VI.

V. Outcomes at Ages Seven and 11

This section examines the impact of Asian flu on test scores, and height at ages seven and 11. As in Section IV, we first establish whether statistically

16. Descriptive statistics of these characteristics are provided in Appendix Table A2. Height is measured in inches and converted to cm.

17. Approximately 11 percent of the NCDS mothers were 60 inches tall or less; 40 percent of mothers smoked prior to pregnancy, evenly divided between heavy and light smokers. Please see Table A1 for summary statistics.

18. There is further variation in epidemic effects by maternal blood pressure and maternal weight: birth weight is reduced where mothers weighed 51 kilograms or less in 1958, or suffered from preeclampsia during pregnancy. However, as both outcomes might be endogenous, results are harder to interpret.

significant effects exist, and test for heterogeneity by maternal health. The panel structure of the data is then used to assess the role of birth weight in capturing or moderating the effects of the fetal health shock.

A. Empirical Method

The baseline specification for estimating the effect of the epidemic upon childhood outcomes, in this example test scores at seven, is as follows:

$$(2) \quad TestScr7_{il} = \alpha + \beta_1 Epid_l + \theta_1 PreEpid_l + \gamma_1 LA_l + \chi_1 X_{il} + \upsilon_{1il}$$

where,

$$(3) \quad \upsilon_{1il} = \tau_{il} g_{il} + \omega_{il} q_{il} + e_l + e_i$$

variables $Epid_l$, $PreEpid_l$, X_{il}, and LA_l are as previously defined. No contemporaneous controls are introduced, as they cannot be assumed independent of the epidemic. The error term, υ_{1il}, is again clustered at the local authority level, and can be decomposed into unobserved genetic and environmental factors, g_{il}, unobserved parental investments, q_{il}, and individual idiosyncratic error terms, e_l and e_i.

$\hat{\beta}_1$ provides a reduced form estimate of the impact of the epidemic upon average child outcomes. However, any significant effects could operate through a number of different channels. These include a direct effect upon health and human capital endowments ($\tau_{il} g_{il}$), or a change in the quantity or efficiency of parental investment ($\omega_{il} q_{il}$).

The role of birth weight in explaining or capturing the effect of the epidemic on childhood outcomes can be examined by comparing Equation 2 to the following specification, which adds birth weight as a control:

$$(4) \quad TestScr7_{il} = \alpha + \beta_2 Epid_l + \mu_2 BW_{il} + \theta_2 PreEpid_l + \gamma_2 LA_l + \chi_2 X_{il} + \upsilon_{2il}$$

The addition of BW_{il} will affect $\hat{\beta}_2$ to the extent that birth weight was a source of correlation between $Epid_l$ and υ_{1il}. As there is typically a positive correlation between child outcomes and birth weight, and the epidemic reduced the birth weight of some cohort members, $\hat{\beta}_2$ should be (weakly) smaller in magnitude than $\hat{\beta}_1$. If any effects of the epidemic are captured entirely by birth weight, $\hat{\beta}_2$ should equal zero.

Irrespective of the relationship between the coefficients estimated by Equations 2 and 4, $\hat{\beta}_2$, should not be interpreted as the causal effect of $Epid_l$, conditional upon birth weight. BW_{il} is what Angrist and Pischke (2009) refer to as a "bad control," as it is itself affected by the epidemic. More specifically, any difference between $\hat{\beta}_1$ and $\hat{\beta}_2$ is consistent with multiple explanations: BW_{il} could act as a second proxy for maternal influenza infection; alternatively, any effects upon test scores may operate through the impact of the epidemic upon birth weight. For the purposes paper, this is unimportant, as the aim is to establish whether birth weight does indeed capture the effect of the fetal shock, and not to provide a causal interpretation for $\hat{\beta}_2$. In subsequent specifications, Equations 2 and 4 are augmented with interactions between $Epid_l$ and maternal height and smoking, replicating the procedure followed in Section IV.

B. Results

Table 3 shows the effects of the epidemic and birth weight upon test scores of cohort members, aged seven and 11. The "Baseline" columns correspond to Equation 2, and the "+BW" columns to Equation 4. The "+Hetero" and "+Hetero and BW" columns augment Equations 2 and 4 with interactions between $Epid_l$, and mother's height and prepregnancy smoking. In contrast to the results for birth weight, the epidemic has a significant negative effect on the mean of both test scores. The "Baseline" specifications indicate implied effects, $\hat{\beta}\left[\frac{sd(Epid_l)}{sd(TestScr_{il})}\right]$, of -0.07 at age seven and -0.06 at age 11, significant at the 1 percent and 5 percent levels, respectively. The $PreEpid_l$ coefficients are positive, as the smaller the difference between $Epid_l$ and $PreEpid_l$, the lower the prevalence of the epidemic.

Estimated coefficients remain statistically unchanged in the "+BW" columns, suggesting that impacts at the mean does not operate through birth weight. Quantile regressions indicate that the epidemic had a statistically significant effect on the middle and upper parts of the conditional test score distributions, between quantiles 0.55 and 0.95 at age seven, and 0.36 and 0.7 at age 11. These results are in keeping with Chay, Guryan, and Mazumder (2009), who find that improved health in early childhood has the greatest effect on test scores at the higher end of the distribution. Birth weight itself has a positive and significant impact on test scores, with implied effects of 0.08 on both test scores, replicating the findings of previous papers to have used the NCDS (for example, Currie and Hyson 1999).[19]

The "+Hetero" columns test for the heterogeneity by maternal health, seen in Section IV. However, three of the four sets of interaction terms in Table 3 are not jointly significant, with high p-values, while the individual interaction terms suggest that the significant heterogeneity by maternal height at seven is spurious.[20] Results do not change when birth weight is included as a control in the "+Hetero+BW" columns; the effect of the epidemic is thus not captured by birth weight, even for groups identified as affected in Section IV. Again, the effect of the epidemic does not vary with any other cohort member or local authority characteristics. There is also no evidence of an association between the epidemic and measures of self or teacher reported parental investment, either as interaction terms added to Equation 4 or when investment measures are used as dependent variables.

The striking differences in the patterns of results for test scores, in Table 3, and birth weight, in Table 2, are consistent with two counterfactual explanations (i) that the fetal health shock impaired cognitive development of some cohort members without affecting observable physical development (ii) that the fetal health shock had a latent effect, which only became apparent as the cohort aged. In Table 4, we assess the impact of the epidemic on a key measure of childhood development, child height (in cm). The specifications in Columns 1 and 3 correspond to Column 4; Columns 2 and 4 add interactions between $Epid_l$ and the two maternal health characteristics.

19. By way of comparison, Rockoff (2004) finds that a one standard deviation increase in teacher quality increases reading and math test scores by 0.1 standard deviations.
20. No interaction is individually significant and the coefficients are not monotonic.

Table 3
Cohort Member Test Scores Aged Seven and 11, the Epidemic and Birth Weight

	Baseline		+BW		+Hetero		+Hetero & BW	
	(1) Age Seven	(2) Age Eleven	(3) Age Seven	(4) Age Eleven	(5) Age Seven	(6) Age Eleven	(7) Age Seven	(8) Age Eleven
$Epid_l$	−0.00152***	−0.00126***	−0.00149***	−0.00124***	−0.00126*	−0.00121**	−0.00128*	−0.00124**
	(0.000548)	(0.000452)	(0.000550)	(0.000457)	(0.000707)	(0.000564)	(0.000708)	(0.000570)
$PreEpid_l$	0.00329**	0.00316**	0.00334**	0.00319**	0.00329**	0.00310*	0.00333**	0.00313**
	(0.00160)	(0.00156)	(0.00159)	(0.00156)	(0.00161)	(0.00158)	(0.00160)	(0.00158)
BW/1000			0.143***	0.157***			0.144***	0.157***
			(0.180)	(0.166)			(0.179)	(0.166)
LA and child-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Maternal smoking × $Epid_l$	No	No	No	No	Yes	Yes	Yes	Yes
Maternal height × $Epid_l$	No	No	No	No	Yes	Yes	Yes	Yes
P-value—no difference in $Epid_l$ by								
Maternal smoking	—	—	—	—	0.5618	0.477	0.6435	0.4049
Maternal eight	—	—	—	—	0.0382	0.6671	0.0308	0.5804
Observations	13,669	12,987	13,669	12,987	13,669	12,987	13,669	12,987
R-squared	0.056	0.149	0.061	0.155	0.056	0.149	0.061	0.155

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent level. The dependent variables normalized Draw a Man Score in Columns 1, 3, 5, and 7, and normalized General Test Score (nonverbal) in Columns 2, 4, 6, and 8. Robust standard errors are clustered at the local authority level. $Epid_l$ represents pneumonia notifications per 100,000 population in LA of birth September-November 1957; $PreEpid_l$ is an average of notifications per 100,000 population in the same periods in 1955 and 1956. Local authority controls come from the 1956 Registrar Generals' returns (rates of still birth and death from TB, and population density) and from the 1951 census (percentage of men leaving school aged 16+, percentage of working age men unskilled, the proportion aged 65 and older, and percentage of households living with >1 person per room). See Table 1 for descriptive statistics. Child-level controls include social class and schooling of both parents, mother's age and its square, the mother's height, the tenure of accommodation and number of persons per room, and whether the mother smoked before pregnancy.

Table 4
Cohort Member Height, in Cm, Aged Seven and 11, Birth Weight and the Epidemic

	+BW		+BW & int	
	(1) Age 7	(2) Age 11	(3) Age 7	(4) Age 11
$Epid_l$	−0.00308	−0.00171	−0.00141	0.00313
	(0.00432)	(0.00330)	(0.00522)	(0.00458)
$PreEpid_l$	0.0130	−0.00495	0.0132	−0.00486
	(0.0132)	(0.0110)	(0.0132)	(0.0113)
BW/1,000	2.073***	2.484***	2.068***	2.475***
	(0.114)	(0.130)	(0.114)	(0.131)
Smoking <10/day × $Epid_l$			−0.00469***	−0.00611***
			(0.00150)	(0.00196)
Smoking 10+/day × $Epid_l$			−0.00620***	−0.00738**
			(0.00201)	(0.00298)
LA and child-level controls	Yes	Yes	Yes	Yes
P-values—joint significance tests for $Epid_l$				
No effect of $Epid_l$	0.4765	0.6050	0.0001	0.0007
No difference by maternal smoking			0.0002	0.0001
No difference by maternal height		0.2810	0.1699	
Observations	12,750	11,679	12,750	11,679
R-squared	0.198	0.199	0.199	0.200

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent level. The dependent variables are the child's height at age 7 (Columns 1 and 3), and child's height at age 11 (Columns 2 and 4), both in cm. Robust standard errors are clustered at the local authority level. $Epid_l$ represents pneumonia notifications per 100,000 population in LA of birth September-November 1957; $PreEpid_l$ is an average of notifications per 100,000 population in the same periods in 1955 and 1956. Local authority controls come from the 1956 Registrar Generals' returns (rates of still birth and death from TB, and population density) and from the 1951 census (percentage of men leaving school aged 16+, percentage of working age men unskilled, the proportion aged 65 and older, and percentage of households living with >1 person per room). See Table 1 for descriptive statistics. Child-level controls include social class and schooling of both parents, mother's age and its square, the mother's height, the tenure of accommodation and number of persons per room, and whether the mother smoked before pregnancy.

The effects of $Epid_l$ upon mean height are very small and not statistically significant: results consistent with (i), but not with (ii). However, interactions added in Columns 2 and 4 indicate significant negative impacts where the mother smoked prior to pregnancy. The same relationship continues when height is measured at 16, and in adulthood. Our results thus suggest that the combination of influenza and smoking can have lasting deleterious effects on physical development. However, the

effects on height do not appear to operate through birth weight, as results are invariant to using BW_{il} as a control.

The differing patterns of results for physical and cognitive development appear consistent with two separate channels of transmission. For physical development, epidemic effects are negative only for certain categories of mothers, where preexisting nutritional stores were possibly low or symptoms plausibly more severe. Effects on cognitive development are invariant to mother or cohort member characteristics, and may stem from inflammation, a characteristic symptom of flu. Given that pandemic flu can sweep through populations of healthy adults and school age children, the heterogeneous effects upon birth weight appears more attributable to differences how a mother's body responds to influenza, rather than to disproportionately higher infection rates. If this is the case, there is no reason to expect that our results for the effects of Asian flu on birth weight and cognitive development should be linked.

VI. Robustness

The discussion in Sections IV and V have assumed that $Epid_l$ is exogenous, conditional on $PreEpid_l$ and LA_l. Yet, three potential confounding factors remain.

First, the epidemic may have generated a nonrandom change cohort composition. Deaths from pneumonia among child-bearing women (15–44) were ten times higher in 1957 than in the year before, but the absolute number remained very small, at just 344 (National Digital Archive of Datasets 2007). Data on still and live births from the Registrar Generals' Returns show that still birth rates were lower in Q1 and Q2 of 1958 than in 1956 or 1957, while live births were higher February, March, and April of 1958, than a year before. Small changes in gestation are impossible to rule out when using monthly data, but the distribution of gestation does not differ markedly from that of the 1970 birth cohort followed by the British Cohort Study (BCS). Furthermore, our results suggest that any effects on gestation were confined to certain subsections of the population. It must also be remembered that the NCDS cohort were exposed between 16 and 25 weeks gestation, after the early stages of pregnancy when spontaneous abortion is most likely. In conclusion, the available aggregate evidence is not suggestive of any substantive change in cohort composition (Nyagu et al. 2002; Otake and Schull 1998).

Second, there data limitations on the number of LA controls that can be constructed for the 1950s. Information from the best available source, the 1951 census, is already included in LA_l. One might therefore be concerned that there remain unobserved LA characteristics, correlated with both excess pneumonia notifications and child outcomes. To address such concerns, we add (ten) region of birth dummies to our preferred specifications for birth weight, Column 4 of Table 2, and test scores, Columns 1 and 2 of Table 3. In each case, results remain statistically unchanged. Confounding factors at the LA level would therefore have to be correlated with $Epid_l$, within region. The potential for bias from unobserved intraregional LA characteristics is still present, but further controls are not available. Comparisons between

the NCDS and the BCS are not possible, as the BCS failed to keep LA of birth records.[21]

Lastly, $PreEpid_l$ and LA_l control only for preexisting LA characteristics and leave any contemporaneous influences, or shocks, unobserved. These shocks enter through ε_{il} or υ_{1il}, and would bias results if correlated with $Epid_l$. The use of region of birth fixed effects goes some way to allaying these fears, as any confounding contemporaneous shock would need to be correlated with $Epid_l$, within region. In the appendix, we present the results using region of birth fixed effects and consider one possible contemporaneous shock, the Windscale Plutonium Factory fire of 1957. Again, there is no evidence that our preferred specifications are biased by other events occurring in the autumn of 1957.

VII. Discussion

This paper finds that Asian influenza did have a significant effect on the childhood outcomes of the NCDS cohort. However, the patterns of results for physical and cognitive development are very different. Impacts of influenza upon physical development, at birth and in childhood, appear contingent upon certain maternal health characteristics: the epidemic reduces birth weight where mothers are short or smoked prior to pregnancy; child height is depressed when mothers smoked. By contrast, effects upon test scores are statistically significant at the mean and middle to upper parts of conditional test score distributions, and are invariant to birth weight, maternal health, socioeconomic status or parental investment. The key implications of our results are as follows.

First, the effects of Asian flu appear to operate through two distinct channels. Physical development growth is impeded where mothers are unable to compensate for interruptions in nutrition, or when symptoms were plausibly more severe. Effects on cognitive development are more general, and may be related to the inflammation that typically accompanies influenza. This demonstrates that responses to a fetal health shock can be heterogeneous. Individual treatment effects, recovered by adjusting estimated coefficients to reflect the average cohort exposure rate, should thus be interpreted with caution.

Second, birth weight does not capture the effect of the epidemic upon child height or test scores. While this finding does not invalidate the existing work that links birth weight to subsequent outcomes, it does demonstrate the need for care when interpreting estimates of "birth weight effects." Birth weight does not represent a catch-all measure of the influences on health and development prior to birth, and the mechanisms that link it to subsequent health and development remain poorly understood. Any biases in existing papers that use the NCDS should be small, as the effect of the epidemic on birth weight is confined to a small group of cohort

21. LA codes for the BCS are only available at ages ten and 16. In our NCDS sample, applying $Epid_l$ dosages for LA of residence at age 11 (1969) produces a null result, which is encouraging as half the sample moved between 1958 and 1969. The absence of statistically significant effects when $Epid_l$ is applied to the BCS, with 1980 LA codes, therefore cannot distinguish between a zero effect of $Epid_l$ and measurement error in the LA of birth.

members, and birth weight has an independent effect on childhood test scores (for example, Case, Fertig, and Paxson 2005; Currie and Hyson 1999).

Third, maternal smoking plays the most consistent and robust role in moderating the effect of the epidemic upon physical growth. Whether these results are attributable to smoking during pregnancy, or to the damage caused by previous smoking behavior, cannot be identified using the NCDS data. The existing public health advice is that women who are pregnant or wishing to conceive should not smoke. Our results perhaps provide one more reason to quit.

Influenza pandemics have occurred three times per century since the 17th century (World Health Organization 2009). During the recent swine flu pandemic, public health campaigns focused upon vulnerable groups, such as young children, the elderly and pregnant women. The results of this paper underline the prudence of this approach, both for both pregnant women, and their unborn children.

Appendix 1

Robustness Results

Table A2 tests for the robustness of our preferred specifications to regional fixed effects and one possible contemporaneous shock. For birth weight, the preferred specification includes interaction terms between the epidemic and maternal height and smoking; for test scores, we use the baseline specifications, without birth weight. Columns labeled "Birth Reg" add region of birth dummies to the preferred specifications. Results will be affected to the extent that these regional variables are correlated with both the outcomes of interest and $Epid_l$. However, for all three outcomes, the estimated coefficients are not statistically different from those in the main text. An unobserved factor at the local authority level, whether preexisting or contemporaneous, could therefore only bias results if it were correlated with $Epid_l$, within region.

The columns labeled "Windscale" add a measure of the log distance from the Windscale plutonium factor, located in the far north west of England. On October 10, 1957, at the peak of the influenza epidemic, a fire started at the Windscale. A significant quantity of radioactive material was emitted before the fire was extinguished on October 12, with Iodine-131 and Plutonium-210 the main particles of concern. The bulk of the emissions were blown down through England, in a southeasterly direction (Crabtree 1959).[22] Given that nuclear radiation can affect fetal

22. The most up-to-date estimates for emissions are as follows: Iodine-131, approximately 1800TBq, with a range of uncertainty of 900–3700TBq; Caesium-137, 90–350TBq; and, Plutonium-210, 42Tbq with a range of 14–110TBq (Garland and Wakeford 2007). Initial public health concerns focused upon Iodine-131, although Plutonium-210 has subsequently been judged more dangerous. A milk ban was put in place in the immediate aftermath of the fire, covering 200 square miles. The concern was that the consumption of milk contaminated by I-131 could lead to thyroid cancer, especially in children. By November 23, the ban had been completely lifted. The incident was rated as a Level 5 accident on the International Nuclear Event Scale. However, emissions of Iodine-131 were 1,000 times less than at Chernobyl some 30 years later. Current estimates place an upper bound on cumulative fatalities of 200, tiny in comparison to an annual death toll from cancer in the U.K. of 100,000 plus (Arnold 1995).

development, there is a concern that the episode could bias our results. However, the estimated $Epid_i$ coefficients remain statistically identical to those in the corresponding "Birth Reg" columns. The same is true when we interact log distance with indicators for direction from Windscale, which take into account the path of the smoke.

Table A1

Summary Statistics of Maternal Characteristics Used to Test for the Heterogeneity of Epidemic Effects across Cohort Members

Variable	Mean	Standard Deviation
Mother smoked before pregnancy		
No	0.59	0.492
Yes—fewer than 10 per day	0.216	0.412
Yes—10 or more per day	0.194	0.396
Mother's height in 1958, in inches (cm)		
$<=60$ ($<=154$cm)	0.109	0.311
61 (154–156 cm)	0.101	0.301
62 (156–159cm)	0.159	0.366
63 (159–161 cm)	0.141	0.348
64 (161–164 cm)	0.163	0.37
65–66 (164–169 cm)	0.18	0.385
$>=67$ (>169cm)	0.106	0.308
Missing	0.04	0.196
N	16,765	

Source: Author's calculations using University of London, IoE (2008a). All characteristics are recorded in 1958. The mean of each dummy variable represents proportion falling into that category dummy

Table A2
Robustness of Preferred Specifications to Regional Fixed Effects and Contemporaneous Shocks

	BW (grams)		Draw a Man Test—Aged 7		General Test—Aged 11	
	(1) Birth Reg	(2) Windscale	(3) Birth Reg	(4) Windscale	(5) Birth Reg	(6) Windscale
$Epid_l$	0.231	0.212	−0.00146***	−0.00154***	−0.00126***	−0.00130***
	(0.293)	(0.293)	(0.000538)	(0.000559)	(0.000421)	(0.000427)
$PreEpid_l$	−0.559	−0.45	0.0016	0.00206	0.00414***	0.00438***
	(0.563)	(0.560)	(0.00173)	(0.00186)	(0.00142)	(0.00148)
$Epid_l$/maternal health interactions						
Smoking—nonsmoker omitted						
$Epid_l \times$ smoking < 10/day	−0.292*	−0.294*				
	(0.160)	(0.159)				
$Epid_l \times$ smoking 10 + /day	−0.413**	−0.415**				
	(0.167)	(0.168)				
Maternal height, inches, 65–66 (164–69cm) omitted						
$Epid_l \times$ < = 60 (< = 154cm)	−0.576***	−0.578***				
	(0.217)	(0.217)				
$Epid_l \times$ 61 (154–56cm)	−0.288	−0.286				
	(0.242)	(0.243)				
$Epid_l \times$ 62 (156–59cm)	0.16	0.163				
	(0.235)	(0.235)				

$Epid_l \times 63$ (159–161cm)	−0.104	−0.107				
l	(0.248)	(0.248)				
$Epid_l \times 64$ (161–164 cm)	−0.132	−0.13				
	(0.243)	(0.242)				
$Epid_l \times >=67$ (>169cm)	−0.28	−0.274				
	(0.582)	(0.581)				
Log distance from Windscale	—	20.65	—	0.0791	—	0.0415
		(22.33)		(0.0626)		(0.0387)
LA and child-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth Region Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,765	16,765	13,669	13,669	12,987	12,987
R-squared	0.084	0.084	0.06	0.06	0.15	0.15

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent level. The dependent variables are birth weight, in grams, in Columns 1 and 2, normalized Draw a Man Score in Columns 3 and 4, and normalized General Test Score (nonverbal) in Columns 5 and 6. Robust standard errors are clustered at the local authority level. $Epid_l$ represents pneumonia notifications per 100,000 population in LA of birth September-November 1957; $PreEpid_l$ is an average of notifications per 100,000 population in the same periods in 1955 and 1956. Local authority controls come from the 1956 Registrar Generals' returns (rates of still birth and death from TB, and population density) and from the 1951 census (percentage of men leaving school aged 16+, percentage of working age men unskilled, the proportion aged 65 and older, and percentage of households living with >1 person per room). See Table 1 for descriptive statistics. Child-level controls include social class and schooling of both parents, mother's age and its square, the mother's height, the tenure of accommodation and number of persons per room, and whether the mother smoked before pregnancy. There are 11 regions of birth: North, North West, East & West Yorkshire, North Midlands, Midlands, East, South East, South, South West, Wales, and Scotland. The omitted region is north. There are a minimum of 10, and a maximum of 32 local authorities per region. The median number of local authorities per region is 14. Log distance from Windscale is measured as distance, in km, from postcode CA20 1PG, to a central postcode in each local authority.

References

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy* 114(4):672–712.

Almond, Douglas, Ken Chay, and David Lee. 2005. "The Costs of Low Birth Weight." *Quarterly Journal of Economics* 120(3):1031–83.

Almond, Douglas, Lena Edlund, Hongbin Li, and Junsen Zhang, J. 2007. "Long-Term Effects of the 1959–1961 China Famine: Mainland China and Hong Kong." NBER Working Paper 13384.

Almond, Douglas, Lena Edlund, and Marten Palme. 2008. "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden." NBER Working Paper 13347.

Almond, Douglas, and Bhashkar Mazumder. 2005. "The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data." American Economic Review: Papers and Proceedings 95(2):258–62.

———. 2008. "The Effects of Maternal Fasting During Ramadan on Birth and Adult Outcomes." NBER Working Paper W14428.

Angrist, Joshua and Jorn-Steffen Pischke. 2009. "Mostly harmless Econometrics: An Empiricist's Companion." Princeton, New Jersey: Princeton University Press.

Arcavi, Lidia and Neil Benowitz. (2004, November). "Cigarette Smoking and Infection." *Archives of Internal Medicine* 164 (20):2206–16.

Arnold, Lorna. 1995. "Windscale 1957. Anatomy of a Nuclear Accident" (2nd ed.). Basingstoke: MacMillan.

Banerjee, Abhijit, Esther Duflo, Gilles Postel-Vinay, and Timothy M. Watts. 2007. "Long Run Health Impacts of Income Shocks: Wine and Phylloxera in 19th Century France." NBER Working Paper 12895.

Barker, David. 1990 "Fetal and Infant Origins of Adult Disease" *British Medical Journal* 301(6761):1111.

Barker, David. 2009. "The Maternal and Fetal Origins of Aging." RAND Mini-Medical School for Social Scientists 2009.

Black, Sandra, Paul Devereux, and Kjells Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122(1):409–39.

Carrat, Fabrice, Elisabeta Vergu, Neil Ferguson, Magli Lemaitre, Simon Cauchemez, Steve Leach, and Alain-Jacques Valleron. 2008. "Time Lines of Infection and Disease in Human Influenza: A Review of Volunteer Challenge Studies." *American Journal of Epidemiology* 167(7):775–85.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24(2):365–89

Cesur, Resul and Inas Rashad. 2008. "High Birth Weight and Cognitive Outcomes" NBER Working Paper No. w14524

Chay, Ken, Jonathan Guryan, and Bhashkar Mazumder. 2009. "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth." NBER Working Paper 15078.

Conley, Dalton and Neil Bennett. 2000. "Is Biology Destiny? Birth Weight and Life Chances." *American Sociological Review* 65(3):458–67.

Crabtree, J. 1959. The Travel and Diffusion of the Radioactive Material Emitted During the Windscale Accident. *Quarterly Journal of the Royal Meteorological Society* 85(366): 362–70.

Currie, Janet, and Rosemary Hyson. 1999. "Is the Impact of Health Problems Cushioned by Socioeconomic Status? The Case of Low Birth Weight." *American Economic Review* 89(2):245–50.
Currie, Janet and Enrico Moretti. 2007. "Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight." *Journal of Labor Economics* 25(2):231–64.
DHHS. 2008. "Flu Pandemics." Department for Health and Human Services, Retrieved 12/3/08, from http://pandemicflu.gov/individualfamily/about/pandemic/index.html
Dos Santos Silva, Isabel, Bianca De Stavola, and Valerie McCormack. 2008. "Birth Size and Breast Cancer Risk: Reanalysis of Individual Participant Data from 32 Studies." *Public Library of Science Medicine* 5(9):e194.
Edwards, Marshall 2007. "Hyperthermia in utero due to Maternal Influenza is an Environmental Risk Factor for Schizophrenia." *Congenital Anomalies* 47:84–89.
Erikson, Willy, John Sundet, and Kristian Tambs. 2009. "Register Data Suggest Lower Intelligence in Men Born the Year After Flu Pandemic." *American Neurological Association* 66:284–89.
Frisbie, E., M. Biegler, P. de Turk, D. Forbers, and S. Pullum. 1997. "Racial and Ethnic Differences in Determinants of Intrauterine Growth Retardation and Other Compromised Birth Outcomes." *American Journal of Public Health* 87(12):1977–83.
Garland, John, and Richard Wakeford. 2007. "Atmospheric emissions from the Windscale Accident of October 1957." *Atmospheric Environment* 41(18):3904–20.
Godfrey, Keith, and David Barker. 2001. "Fetal Programming and Adult Health." Public *Health Nutrition* 4(2B):611–24.
Goldenberg, Robert, and Dwight Rouse. 1998. "Prevention of Premature Birth." New *England Journal of Medicine* 339(5):313–20.
Goodnight, William, and D. E. Soper. 2005. "Pneumonia in Pregnancy." *Critical Care Medicine* 33(10):390–97.
Guerri, Consuelo, Alissa Bazinet, and Edward Riley. 2009. "Foetal Alcohol Spectrum Disorders and Alterations in Brain And Behavior." *Alcohol and Alcoholism* 44(2):108–14.
Hunter, John, and Jonathan Young. 1971. "Diffusion of Influenza in England and Wales." *Annals of the Association of American Geographers* 61(4):637–53.
Jaddoe, Vincent, and Jacqueline Witteman, J. 2006. "Hypotheses on the Fetal Origins of Adult Diseases: Contributions of Epidemiological Studies." *European Journal of Epidemiology* 21(2):91–102.
Kramer, Michael. 1987. "Intrauterine Growth and Gestational Duration Determinants." *Pedriatics* 80(4):502–11.
Lambert, Scott. 2007. "Congenital Rubella Syndrome: the End is in Sight." British *Journal of Ophthalmology* 91:1418–19.
Ling, Min-Jen. 2008. "Does "in utero" Pandemic Exposure Matter? Evidence from the Long Term Effect of the 1918 Influenza on Educational Attainments in Taiwan." National Taiwan University Working Paper.
Martorell, Reynaldo and Teresa Gonzalez-Cossio, T. 1987. "Maternal Nutrition and Birth Weight." *Yearbook of Physical Anthropology* 30:195–220.
Metzger, Boyd, Veronica Ravnika, Rita Vileisis, and Norbert Freinkel. 1982. "Accelerated Starvation and the Skipped Breakfast in Late Normal Pregnancy." *Lancet* 8272(1):91–102.
Ministry of Health. 1960. "Influenza epidemic in England and Wales, 1957–58" Reports on Public Health and Medical Subjects (100).
National Digital Archive of Datasets. 2007. "Historical mortality files—CRDA/20." National Digital Data archives. Retrieved 29/1/08, http://www.ndad.nationalarchives.gov.uk/CRDA/20/detail.html.

Nelson, Richard. 2010. "Testing the Fetal Origins Hypothesis in a Developing Country: Evidence from the 1918 Influenza Pandemic." *Health Economics* 19(10):1181–92.
NHS. 2009. "Smoking and Pregnancy: Go Smokefree—for You and Your Baby." National Health Service. Retrieved 09/21/09, from http://smokefree.nhs.uk/smoking-and-pregnancy/
Nilsson, Peter. 2008. "Does a Pint a Day Affect your Child's Play? The Effect of Prenatal Alcohol Exposure on Adult Outcomes." Cemmap Working Paper CWP22/08
Nyagu, Angelin, Konstantin Loganovsky, Tatiana Loganovskaja, Viktor Repin, and Stanislave Nechaev. 2002. "Intelligence and Brain Damage in Children Acutely Irradiated in Utero As a Result of the Chernobyl Accident." *KURRI KR* 79:202–30.
O'Callaghan, E., P. Sham, N.Takei, G. Glover, and R. Murray. 1991. "Schizophrenia After Prenatal Exposure to 1957 A2 Influenza Epidemic." *Lancet* 337 (8752):1248–50.
Otake, Masanori and William Schull. 1998. "Review: Radiation-Related Brain Damage and Growth Retardation Among the Prenatally Exposed Atomic Bomb Survivors." *International Journal of Radiation Biology* 74(2):159–71.
Rasmussen, Sonja, Denise Jamieson, and Joseph Bresee. 2008. "Pandemic Influenza and Pregnant Women." *Emerging Infectious Diseases* 14(1):95–100.
Registrar General for England and Wales. 1957. "The Registrar General's Weekly Return for England and Wales. Births and Deaths, Infectious Diseases, and Weather." HMSO, London.
Registrar General for Scotland. 1957. "Weekly Return of the Registrar General Scotland. Births, Deaths, and Marriages. Infectious Diseases. Weather report" HMSO, Edinburgh.
Rockoff, Jonah. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94(2):247–52.
Scrimshaw, Nevin. 1977. "Effect of Infection on Nutrient Requirements." *American Journal of Clinical Nutrition* 30:1536–44.
Simonsen, Lone, Matthew Clarke, Lawrence Schonberger, Nancy Arden, Nancy Cox, and Keiji Fukuda. 1998. "Pandemic versus Epidemic Influenza Mortality: A Pattern of Changing Age Distribution." *Journal of Infectious Diseases* 178:53–60.
Tomkins, Andrew, Susan Murray, Patricia Rondo, and Suzanne Filteau. 1994. "Impact of Maternal Infection on Fetal Growth and Nutrition." Retrieved 12/01/2008, from http://www.unsystem.org/scn/archives/scnnews11/ch08.htm.
University of London. IoE. 2008a. National child Development Study: Childhood Data, Sweeps 0–3, 1958–74 (SN5565). Computer file, Distributed by UKDA.
———. 2008b. "National Child Development Study: Local Authority Data, 1958–1974: Special Licence Access." Computer File, Distributed by UKDA.
Woodall, John, K. E. K. Rowson, and J.C. McDonald. (1958). "Age and Asian influenza, 1957." British Medical Journal, 2 (5108): 1305–1368.
World Health Organization. 2009. Retrieved 2009-01-26, from http://www.who.int/csr/disease/influenza/pandemic/en/