Identifying the causal impact of poor in-utero health on later-life outcomes is difficult given that unobservables which determine these outcomes also predict the probability of poor in-utero health. In order to combat this, previous studies have made use of natural experiments, such as pandemics, which generate exogenous variation in the fetal environment. However, data on in-utero health at an individual level is rarely available, making estimation of a treatment effect on the treated difficult. In addition, studies examining effects of in-utero health on mental health outcomes remain scarce. This represents an important gap in the literature, given that associations between mental health outcomes in childhood and health and labour market outcomes in adulthood are well-established. We build on the previous literature by examining the effects of in-utero exposure to influenza on mental health in childhood and survival at various ages, using the outbreak of the 1957 Asian influenza epidemic in the UK as a natural experiment. We use data from the National Child Development Study (NCDS), whose cohort members were exposed to the epidemic prior to birth. In contrast to many datasets, the NCDS provides individual-level data on in-utero exposure to influenza, derived from a question asking mothers of cohort members to self-report whether they contracted influenza during pregnancy. Results suggest that in-utero exposure to influenza has little effect on mental health in childhood, but significantly reduces age at death, primarily through an increase in the probability of being stillborn. To deal with potential endogeneity, we employ an additional identification strategy, involving the use of plausibly exogenous geographical variation in the intensity of the epidemic as an instrument for individual-level exposure to influenza in-utero. However, this area-level instrument provides only weak identification of individual-level exposure, raising a potential issue regarding an identification strategy used pervasively in studies examining the effects of in-utero health on later-life outcomes.

Acknowledgements: We are grateful to Elaine Kelly for providing local authority-level data on influenza intensity.
1. INTRODUCTION

Half of mental disorders in adults begin before the mid-teenage years (Kessler et al., 2007; World Health Organization, 2013). This is corroborated in the academic literature, where poor mental health in childhood has been found to be associated with the incidence of mental health problems at various stages of adulthood (Carneiro et al., 2007; Jones et al., 2011; Kaestner and Callison, 2011). Given these associations, a well-documented upward trend in child and adolescent mental health problems is concerning (Bricker et al., 2004; Hagell et al., 2013; World Health Organization, 2003). However, the consequences of childhood mental health problems look even more severe once a broader definition of mental health is considered. Social and emotional well-being serves as an umbrella for a wide variety of wellbeing measures, covering social wellbeing (measuring the quality and existence of relationships with others), psychological well-being (such as the ability to manage emotions and be empathetic) and emotional well-being (including being confident, happy and not anxious and depressed). It also covers a range of non-cognitive skills and personality traits studied heavily in the economics literature, including self-discipline (Duckworth and Seligman, 2005), self-motivation (Heckman et al., 2006), social adjustment (Carneiro et al., 2007; Goodman et al., 2011; Jones et al., 2011), self-esteem (Blanden et al., 2007; Heckman et al., 2006; Kaestner and Callison, 2011; Murasko, 2007), locus of control (Blanden et al., 2007; Murasko, 2007), and self-regulation (Conti and Heckman, 2010).

Under this definition, findings from US and UK birth cohorts suggest that mental health problems in childhood are not only associated with adult mental health problems but also a range of other important outcomes throughout the life-course. In a health setting, there is evidence to suggest that poor social and emotional well-being in childhood has a lasting effect on adult physical health and health behaviours, covering measures such as the likelihood of smoking (Carneiro et al., 2007; Conti and Heckman, 2010), obesity (Conti and Heckman, 2010), disability and long-standing illness (Jones et al., 2011), self-assessed health (Murasko, 2007), and premature mortality (Jokela et al., 2009; Maughan et al., 2013; Turner et al., 2014). Poor childhood social and emotional well-being also has important implications for the labour market, with associations found for earnings and employment (Blanden et al., 2007; Carneiro et al., 2007; Currie et al., 2010; Fletcher, 2013; Heckman et al., 2006; Johar and Truong, 2014; Smith and Smith, 2010) and measures of human capital development such as the level of education (Carneiro et al., 2007; Currie et al., 2010; Duckworth and Seligman, 2005; Fletcher and Wolfe, 2008; Fletcher, 2010; Heckman, 2006; Smith and Smith, 2010; Washbrook et al., 2013). In some cases, the association between non-cognitive skills and labour market outcomes are of a similar (and often larger) magnitude than for cognitive skills, which were traditionally seen as the key driver of success (Carneiro et al., 2007; Duckworth and Seligman, 2005). Effects also extend to social outcomes, with poor social and emotional well-being found to be associated with increases in the likelihood of criminal activity (Carneiro et al., 2007) and a reduction in the probability of marriage (Goodman et al., 2011; Smith and Smith, 2010).

Given the severe negative consequences of poor mental health in childhood and social and emotional well-being more widely, it is important that its determinants are known. Numerous studies have investigated the determinants of poor mental health, with many finding that the probability of onset is related to a range of social, economic and environmental factors (Allen et al., 2014; World Health Organization, 2014). Recent studies in neuroscience, molecular biology, epigenetics, and the behavioural sciences hypothesise that this may be due to adverse events in childhood stimulating a stress response, causing brain circuits to be impaired and thereby increasing the likelihood of long-term problems in mental health (Shonkoff, 2011).

A growing body of literature in both economics and public health have recognised the important role pre-natal health can play in determining later-life outcomes. This literature is based primarily on the fetal origins hypothesis. In its original form, this hypothesis argues that inadequate nutrition in-utero leads to increased risk of the development of disease in adulthood (Barker,
1990). Economists have since extended the fetal origins hypothesis beyond nutrition, through the study of various health behaviours and environmental factors which may influence fetal well-being (Almond and Currie, 2011). One suggested mechanism is that fetal under-nutrition, or other fetal exposures, “reprogram” an individual’s epigenome, setting “switches” which determine which parts of the genome are expressed, and which consequently determines the likelihood of disease onset (Petronis, 2010). Models developed by James Heckman and colleagues have formalised the process through which in-utero conditions may impact later-life outcomes (Cunha and Heckman, 2008; Heckman, 2007). These models suggest that human capacity is a multi-dimensional concept, encompassing aspects of health, cognitive skills, and non-cognitive skills. They model capacity as a dynamic process using the ideas of “self-productivity”, where high levels of capacity in one period results in higher capacity in all future periods, and “dynamic complementarity”, where the productivity of investments to improve capacity in one period is increasing in the level of capacity in the previous period. Good in-utero health can therefore be treated as an investment which improves capacity both contemporaneously and throughout the life-course.

Tests of the fetal origins hypothesis and its extensions are plagued by the existence of omitted variables which are correlated both with the likelihood of poor in-utero health and the incidence of poor later-life outcomes. This non-random selection into poor in-utero health means that the identification of causal effects is difficult. As a result, studies in the economics literature have often relied on natural experiments to generate exogenous variation in fetal health conditions, in order to identify causality. Examples of these natural experiments include famines (Almond et al., 2010; Banerjee et al., 2010; Scholte et al., 2012), religious fasting (Almond et al., 2014), pollution and radioactive fallout (Almond et al., 2009; Bharadwaj et al., 2014; Black et al., 2013; Currie and Walker, 2011; Isen et al., 2014; Otake and Schull, 1998; Sanders, 2012), and changes in alcohol consumption in response to policy (Nilsson, 2008). These have been supplemented by the more recent literature studying the lifecycle consequences of perinatal interventions (Bhalotra et al., 2015a, 2015b; Maselko et al., 2015).

One of the most commonly-used natural experiments in the economics literature has been influenza pandemics, with identification strategies making use of both temporal and geographical variation in pandemic intensity. A substantial body of research has provided convincing evidence that in-utero exposure to the 1918 Influenza pandemic, which disproportionately affected pregnant women, resulted in deleterious effects on many adult outcomes. These include health outcomes such as diabetes (Almond and Mazumder, 2005; Garthwaite, 2008; Lin and Liu, 2014), cardiovascular disease (Garthwaite, 2008; Mazumder et al., 2010), kidney disease (Garthwaite, 2008; Lin and Liu, 2014), respiratory problems (Lin and Liu, 2014), and the incidence of strokes (Almond and Mazumder, 2005); and labour market and human capital outcomes such as educational attainment (Almond, 2006; Garthwaite, 2008; Lin and Liu, 2014; Neelsen and Stratmann, 2012; Nelson, 2010; Richter and Robling, 2013), literacy (Nelson, 2010), income (Almond, 2006; Nelson, 2010) and employment (Nelson, 2010).

Despite the abundance of evidence regarding its effects on adult outcomes, with the exception of birthweight1, studies examining effects of poor in-utero health on outcomes in childhood are relatively few in number. In addition, data restrictions mean that studies in a UK setting are rare. Kelly (2011) represents an exception to the above. They use data from the National Child Development Study (NCDS), to study the effects of in-utero exposure to influenza on physical health and cognitive ability in childhood. The NCDS is a British cohort dataset following a group of approximately 17,000 individuals born from 2nd-9th March 1958 from birth until present day, collecting data throughout childhood and adulthood. Their identification strategy uses exogenous geographical variation in the severity of the 1957 Asian influenza epidemic across 172 Local Authorities of Great Britain. They find that in-utero exposure to influenza led to a reduction in birthweight and reductions in physical development.

---

1 See Douglas Almond & Currie (2011) and J Currie & Rossin-Slater (2015) for reviews.
(measured using height) and cognitive ability at ages 7 and 11. They suggest that their results for physical development are driven by maternal influenza causing a reduction in fetal nutrition, and that findings for cognitive ability are as a result of maternal influenza hindering brain development, which is at its most important stages between 8 and 25 weeks into gestation, the period in which the 1957 epidemic was at its strongest.

We use a similar identification strategy to Kelly (2011) and again study the effects of fetal exposure to influenza on child development at age 7 and 11. However, we make several additional contributions to their approach, and to the wider fetal origins literature. Firstly, we study the impacts of fetal exposure to influenza on mental health, a hugely understudied outcome in the fetal origins literature (Adhvaryu et al., 2015). This study also represents one of the first studies to examine mental health at childhood ages. Given its associations with later-life outcomes, this is a substantial gap in the literature. Secondly, by linking the NCDS to health service death records, we also study the effects of influenza exposure on the longevity of this same cohort, a similarly understudied outcome (Bhalotra et al., 2015a). Thirdly, as there are several hypothesised pathways between childhood mental health problems and early mortality (Angold, 2009), and some evidence of mortality effects (Jokela et al., 2009; Maughan et al., 2013), we investigate the proportion of the effect of in-utero exposure to influenza on longevity that can be explained by its effects on childhood mental health. Fourthly, we use a strengthened identification strategy that makes use of newly-released mother-level data on influenza exposure during pregnancy. Instead of including geographical influenza rates directly in the regression specification, this new strategy uses the plausibly exogenous geographical influenza rate variation as an instrument for child-level fetal exposure. As a result, we can estimate the local average treatment effect, or the treatment effect on the treated, rather than just the average treatment effect. In addition, in a final and possibly minor contribution, we also control for a richer set of background characteristics than Kelly (2011).

We find that in-utero exposure to influenza has little impact on mental health outcomes in childhood, but that it generates a significant reduction in survival rates, primarily driven by effects on the rate of stillbirths. Given limited effects on mortality in the later stages of the life-course, mental health in childhood plays a limited role in explaining influenza’s effects on longevity.

The remainder of this paper is organised as follows. Section 2 provides a description of the 1957 Asian influenza epidemic and its regional variation within the UK. Section 3 outlines how fetal exposure to influenza could affect the non-cognitive development and mental health of children, and the mechanisms through which mental health in childhood could reduce life expectancy. The empirical strategy is then described in Section 4. Section 5 outlines the data, before the results are presented in Section 6. The paper concludes in Section 7 with a discussion of results, the limitations of the paper, plans for future work and some implications for policy makers.

## 2. THE 1957 ASIAN INFLUENZA EPIDEMIC

The Asian Flu epidemic hit Great Britain through its seaports in early June, 1957. The outbreak reached epidemic levels in September 1957 and lasted for 12 weeks until dissipating in November 1957 (Hunter and Young, 1971). Although estimates vary, it was approximated that the epidemic led directly to between 6 and 9 million influenza cases (13-19% of the population) and to between 6,000 and 14,000 deaths (Epidemic Observation Unit, 1958; Hunter and Young, 1971). The epidemic primarily affected children and young adults, with the highest amount of cases occurring in those aged 4-15. One study even suggested that the epidemic may have been almost entirely confined to school-children and their families (Woodall et al., 1958). Pregnant women are particularly susceptible to significant morbidity as a result of influenza (Hunter and Young, 1971). Statistics from the
The epidemic appear to corroborate this, with the influenza rate being approximately 30% for women of childbearing age (Woodall et al., 1958), although separate statistics for pregnant women are not available.

The primary focus of the epidemic was in Northern England, but quickly spread to the South, East and Scotland in mid-September (Jackson, 2009), with significant within-region variation (Hunter and Young, 1971). The spread of the epidemic was shown to follow lines of high population potential (Hunter and Young, 1971).

The epidemic peaked in the week ending 17th October 1957 (Jackson, 2009). Given that the NCDS follows those born 2nd-9th March 1958, this means 90% of the NCDS children were between 17 and 23 weeks in gestation in this peak period.

3. MECHANISMS THROUGH WHICH INFLUENZA COULD AFFECT MENTAL HEALTH AND LONGEVITY

3.1. EFFECT OF INFLUENZA ON CHILD MENTAL HEALTH

The primary mechanism through which in-utero exposure to influenza could affect mental health in childhood is through a direct effect on fetal brain development. In-utero development of the brain occurs between 8 and 25 weeks into gestation, and in this period the brain is particularly vulnerable to adverse fetal conditions (Nyagu et al., 2002).

Development of the brain occurs in two periods. In the first of these, occurring 8 to 15 weeks into gestation, rates of proliferation of nerve cells (neural elements) and migration of neurons to different parts of the developing brain are at its highest. Previous studies have found evidence of the most severe effects of fetal health shocks on cognitive development in this period (Otake and Schull, 1998). Given effects on cognitive ability in this period, in-utero exposure to influenza in this period is also likely to have an effect on the development of non-cognitive skills, a key aspect of mental health in childhood. The second stage of development occurs between 16 and 25 weeks into gestation. This period is when neurons develop their biochemical and physiological properties, the creation of new synapses (synaptogenesis) hits its peak, the brain’s architecture begins to form, and neuronal pruning takes place (Nyagu et al., 2002). There is existing, although not causal, evidence that in-utero exposure to influenza during the Asian Flu epidemic in this second development period interrupted the neural pruning process, leading to schizophrenia in adulthood (O’Callaghan et al., 1991). The epidemic hit its peak in the week ending the 17th of October 1957, when 95% of the NCDS cohort was between 16 to 25 weeks into gestation, and so the majority of the in-utero exposures in this cohort would have occurred in the second period of brain development (Kelly, 2011). Exposures to other cohort members would have occurred in the first period of brain development, between 8 and 15 weeks into gestation.

However, this previous evidence does not provide exact mechanisms through which influenza could affect child mental health.

Figure 1 outlines a model, taken from Schlotz & Phillips (2009), which suggests how a range of maternal health conditions and behaviours could affect the fetal environment and how this could subsequently affect mental health in childhood and beyond. Maternal health and behaviours, such as nutrition, smoking, alcohol consumption, psychosocial stress and infection, can affect the fetal environment both indirectly and directly. The indirect pathways occur through changes in the levels of oxygen and micronutrients supplied to the fetus caused by the actions of stress hormones. The direct pathway occurs through the transfer of maternal nutrients across the placenta. These then effect brain development, neurotransmitter systems, and neuroendocrine systems, which determine the child’s behaviour and mental health once born.

Many pathways described within this model are potentially relevant when examining the effect of in-utero exposure to influenza. As outlined briefly above, evidence from the psychiatry literature suggests that in-utero exposure to maternal infection substantially increases the risk of schizophrenia emerging in adulthood, with some evidence specifically related to
influenza (Boksa, 2008; Brown, 2006; O’Callaghan et al., 1991). Possible mechanisms for this include hypothermia and inflammation interrupting the neural pruning process during the brains development (Rasmussen et al., 2008).

3.1.1. NUTRITION

A fetus receives nutrients from its mother via the placenta from both her diet and the stock of nutrients stored in her liver. Kelly (2011) suggests that maternal influenza could affect fetal nutrition through multiple mechanisms. Most importantly, influenza acts to suppress appetite, leading to substantially reduced nutritional intake. Evidence suggests that higher levels of micronutrients such as iron and fatty acids as well as fish intake during pregnancy could lead to reductions in behavioural problems in later-life (Colombo et al., 2004; Gale et al., 2008; Parsons et al., 2008). This problem is compounded by influenza increasing excretion rates and interfering with the absorption of nutrients (Tomkins et al., 1994).

3.1.2. MATERNAL PSYCHOSOCIAL STRESS

The contraction of influenza during pregnancy could also increase maternal stress. A growing body of evidence suggests that increased stress during pregnancy increases the risk of the offspring developing behavioural difficulties and a range of mental disorders such as depression, schizophrenia and autism (see Talge, Neal, & Glover (2007) for review). Mechanisms for this include stress disrupting both the functioning of hormone-producing glands (Weinstock, 2008) and the protective quality of the placenta (Schmitt et al., 2014).

3.1.3. COMMON MECHANISMS

Previous literature suggests different prenatal factors such as maternal smoking and stress have remarkably similar effects on mental health outcomes, suggesting that these effects are likely driven by similar mechanisms affecting neurodevelopmental changes. Schlotz & Phillips (2009) highlight fetal programming as one potential mechanism. Hyperactivity/inattention during childhood may be a “response ready” trait that develops as a response to a resource-depleted or fast-changing fetal environment (Jensen et al., 1997). Development of this trait is an evolutionary response as a result of foreseeing a similarly fast-changing post-natal environment. This trait then turns out to be a disadvantage when the post-natal environment is not as fast-changing as in-utero.

3.1.4. ECONOMIC CONDITIONS

News reports published at the time of the 1957 epidemic suggest that areas in which the intensity of the outbreak was high could have experienced a substantial deterioration in economic conditions (Jackson, 2009). In Manchester, where cases of influenza rose sharply in the epidemic period, factories, offices and mines closed, leading to the Manchester Guardian publishing a headline “Setback in Production – ‘Recession through Influenza’” (Manchester Guardian, 29th November 1957). Parental job loss and/or reductions in income would have led to a reduction in the resources available for parental investments in their children, an important determinant of childhood health (Heckman, 2007). A reduction (or a fear of a reduction) in economic fortunes could also increase maternal stress, causing a further deterioration in fetal health.

3.2. EFFECT OF INFLUENZA ON LONGEVITY AND THE INDIRECT EFFECTS OF POOR CHILD MENTAL HEALTH

As mentioned previously, the fetal origins hypothesis predicts that in-utero exposure to health shocks leads to the fetus being programmed with particular metabolic traits that increase the likelihood of disease in adulthood (Barker, 1990). One possible explanation is that fetal health shocks set “switches” in the epigenome, which cause certain genes to be expressed later in life
(Petronis, 2010). If we make the plausible assumption that those not experiencing disease are likely to live longer than those that do, then by assumption the fetal origins hypothesis also predicts that fetal health shocks will increase the likelihood of premature mortality.

However, the fetal origins hypothesis assumes that the metabolic traits programmed in-utero remain latent until late-adulthood, and so would not predict a mortality differential in the earlier parts of the life-course. However, models of skill formation developed by James Heckman and colleagues suggest otherwise. They model skills, health included, as a dynamic process in which ‘skill begets skill’ such that higher levels of health in one period would increase health in all future periods (Cunha and Heckman, 2007; Heckman, 2007). As a result, a negative health shock in-utero could reduce health, and thus cause an increase in the probability of death, even in the early years. In addition, evidence from natural experiments suggest that fetal health shocks reduce the level of education and earnings in the labour market (Almond and Currie, 2011), which may further reduce the level of health, independent of these metabolic traits (Cutler et al., 2006; Cutler and Lleras-Muney, 2010).

In addition to these direct effects on longevity, it is likely that poor mental health in childhood represents an important pathway through which fetal health affects premature mortality. Drawing on evidence from a vast literature studying the associations between childhood mental health, markers of health, and causes of death at different stages of the life-course, Angold (2009) suggest a number of potential mechanisms through which poor mental health in childhood could increase the likelihood of premature mortality, including higher rates of substance abuse, a higher risk of suicides, and a higher prevalence of physical conditions related to psychiatric disorders.

Given these mechanisms, if the effects of in-utero exposure to influenza on mental health outcomes are strong, they are likely to explain a substantial portion of the mortality effects, and do so at many ages throughout the life-course.

4. EMPIRICAL STRATEGY

4.1. BASIC SPECIFICATION

In the first step, we do not exploit data from the epidemic, and estimate the following regression specification:

\[ Y_{itra} = \alpha + \beta f\flu_{itra} + \rho X_{itra} + \gamma LA_t + \theta_r + \epsilon_{itra} \]  

(1)

Here \( Y_{itra} \) represents age \( a \) outcomes (mental health, \( MH_{itra} \), and/or mortality, \( D_{itra} \)) for child \( i \) born in local authority \( l \). \( f\flu_{itra} \) is a binary variable, which equals unity if child \( i \) is exposed to maternal influenza in-utero, and zero otherwise. \( X_{itra} \) is a vector of pre-determined child-level controls, including child characteristics and measures of family background, which are likely to predict both the likelihood of in-utero exposure to influenza and the outcomes of interest. \( LA_t \) is a vector of local authority characteristics, \( \theta_r \) is a set of fixed effects for each of Britain’s ten regions, and \( \epsilon_{itra} \) is a zero-mean error term assumed uncorrelated with all other variables in the model. Equation (1) is estimated using OLS, and standard errors are clustered at the local authority level.

However, the estimate of \( \beta \), the effect of in-utero exposure to influenza on outcomes, in equation (1) is likely to be biased. Unobserved factors could be correlated with both \( f\flu_{itra} \) and \( Y_{itra} \), resulting in \( E[\epsilon_{itra}|f\flu_{itra} \neq 0] \). An example of this could be genetic endowments of poor health shared between mothers and their offspring, the first increasing the likelihood of

---

2 For a review of evidence from delinquent samples of young men see Piquero et al., (2014). For results from population samples see Maughan et al., (2013), Jokela et al. (2009), and Colman et al. (2009).
contracting influenza, the latter increasing the likelihood of poor mental health in childhood and premature mortality. A second example would be measurement error. An over/under-reporting of influenza might be due to unobserved maternal characteristics that then affect the outcomes of child $i$.

4.2. INSTRUMENTAL VARIABLE APPROACH

In order to circumvent the endogeneity issue above, we estimate the following regression specification:

$$f\text{lu}_{ilr} = \delta + \pi X_{ilr} + \sigma LA_{l} + \varphi PreEpid_{l} + \tau Epid_{l} + \theta_{r} + u_{ilr} \quad (2)$$

$$Y_{ira} = \alpha + \beta f\text{lu}_{ilr} + \rho X_{ilr} + \gamma LA_{l} + \varphi PreEpid_{l} + \theta_{r} + \epsilon_{ira} \quad (3)$$

Here $Epid_{l}$ represents the influenza infection rate in local authority $l$ in the epidemic period, and $PreEpid_{l}$ represents the influenza infection rate in local authority $l$ in the period prior to the epidemic. We allow pre-epidemic influenza rates to directly influence outcomes, and this will capture any correlation between the underlying health of local authority populations and the prevalence of influenza. The local authority influenza rates in the epidemic period thus serve as our exclusion restriction. The instrument can therefore be interpreted as the epidemic-induced deviations in infections away from their long-term trends at the local authority level. These deviations are likely to be highly correlated with individual-level influenza exposure and so are likely to represent strong instruments. Equations (2) and (3) are estimated using two-stage least squares, and errors are again clustered at the local authority level. The error terms of the first and second stage equations, $u_{ilr}$ and $\epsilon_{ira}$, are assumed to be uncorrelated.

The key identifying assumption here is that these deviations are exogenous with respect to $\epsilon_{ira}$. The inclusion of local authority level controls in this case is important given that they may be correlated with both outcomes and disease spread. The inclusion of region fixed effects control for the effects of any unobserved local authority characteristics which are fixed within region. We therefore restrict bias to be caused by unobserved local authority-level confounders which vary over time and within region.

4.3. ESTIMATING THE ROLE OF CHILDHOOD MENTAL HEALTH IN EXPLAINING THE EPIDEMICS EFFECTS ON LONGEVITY

We estimate the following regression specification:

$$f\text{lu}_{ilr} = \delta + \pi X_{ilr} + \sigma LA_{l} + \varphi PreEpid_{l} + \tau Epid_{l} + \theta_{r} + u_{ilr} \quad (4)$$

$$D_{ira} = \alpha + \beta f\text{lu}_{ilr} + \rho X_{ilr} + \gamma LA_{l} + \varphi PreEpid_{l} + \omega MH_{ilr} + \theta_{r} + \epsilon_{ira} \quad (5)$$

$D_{ira}$ represents a binary variable which equals one if individual $i$ is dead by age $a$, and $MH_{ilr}$ denotes the mental health outcomes of individual $i$ in childhood. Here, we must restrict the sample to those who survive past childhood. The role of childhood mental health in explaining the effect of in-utero exposure to influenza on longevity can be tested by comparing $\hat{\beta}$ from equation (5) to $\hat{\beta}$ in equation (3), when estimated on the same sample. As we expect exposure to influenza to increase the likelihood of death and reduce mental health, we expect the inclusion of $MH_{ilr}$ to shift $\hat{\beta}$ towards zero. If the effects of influenza exposure on longevity is captured entirely by its effects on childhood mental health, then $\hat{\beta}$ in equation (5) should equal zero.

As stated in Kelly (2011), the coefficients of a variable included on the causal pathway between two other variables cannot be interpreted as causal. In this case, mental health would be classed as a bad control, as it is itself an outcome of the epidemic (Angrist and Pischke, 2008). Identification of this effect would require an additional instrument for mental health. However, the
aim here is not to identify the causal effect of mental health on mortality, only to explain whether mental health captures the mortality effects of in-utero exposure to influenza.

5. DATA

5.1. THE NATIONAL CHILD DEVELOPMENT STUDY

Our primary data source is the National Child Development Study (NCDS). The NCDS is a longitudinal study which follows a cohort of approximately 17,000 individuals born in England, Scotland and Wales between the 2nd and 9th of March, 1958. Initially funded by the National Birthday Trust Fund, the NCDS began with the Perinatal Mortality Survey (PMS), which collected data via questionnaires completed by mothers and midwives in attendance at delivery, on factors associated with stillbirth and infant death. Since then there have been nine subsequent follow-up interviews aimed at gathering information on a range of health, education, social and economic outcomes. These follow-up interviews took place at ages 7 (1965), 11 (1969), 16 (1974), 23 (1981), 33 (1991), 42 (1999/2000), 46 (2004/05), 50 (2008/09) and 55 (2013/14). Data at these follow-ups was drawn from a variety of sources including the cohort members themselves, their parents and partners, local authority medical officers, and schools. 17,415 individuals completed the PMS, with this falling to 15,425 in the age 7 survey and 15,337 in the age 11 survey. Completion rates continued to fall as the cohort aged, with only 9,137 completing the age 55 survey.

5.1.1. CHILD-LEVEL EXPOSURE TO INFLUENZA

Our main variable of interest is in-utero exposure to influenza. This is derived from the PMS, in which NCDS mothers were asked “Were any of the following abnormalities or illnesses, or any other condition, encountered in pregnancy?” of which influenza was one of the conditions listed. Due to poor coding, in earlier releases of the NCDS influenza cases could not be separated from other conditions in pregnancy, explaining why previous examinations of the effect of in-utero influenza in the NCDS cohort failed to exploit data on individual-level exposure (Kelly, 2011). However this data has been available since 2014 as a result of a conversion process conducted by the Centre for Longitudinal Studies, the current organisers of the NCDS (Dodgeon et al., 2014).

5.1.2. CHILDHOOD MENTAL HEALTH

Our primary outcomes of interest are two measures of childhood mental health at ages 7 and 11. The first of these is derived from the Rutter Behaviour Scale (Rutter et al., 1970), which is the pre-cursor the Strengths and Difficulties Questionnaire, widely regarded as one of the gold standard instruments for the measurement of social and emotional well-being in children (Wolpert et al., 2009). When cohort members were aged 7 and 11, parents of each cohort member were asked to state the frequency that they observed their child engaging in range of negative behaviours on a scale of “Never”, “Sometimes”, “Frequently”. In the complete scale, 31 such behaviours are included, but only 14 are used in the NCDS (Centre for Longitudinal Studies, 2012). Responses are recorded as 0, 1 and 2 respectively, and are summed to create an overall index ranging from 0 (perfect mental health) to 28 (worst mental health). Some parents fail to complete a subsection of the Rutter questions. If responses to 4 or more questions are missing, a missing value is assigned to the index. For those with less than 4 missing responses, missing values for these responses were replaced with the average response to the non-missing responses. An indicator variable which equals

---

3 The others were diabetes, heart disease, tuberculosis, German measles, vaginal bleeding <28 weeks in gestation, suspected disproportion, psychiatric disorders (under treatment), external version, and three types of bleeding in pregnancy.

4 In the NCDS, parents are asked whether their child: “Is squirmir or fidgety”, “Destroys own or others’ belongings [e.g. tears or breaks]”, “Fights with other children”, “Worries about many things”, “Prefers to do things on his/her own rather than with others”, “Is irritable, often flies off the handle”, “Is miserable or tearful”, “Has twitches or mannerisms of the face, eyes or body”, “Sucks thumb or finger during the day”, “Bites nails”, “Is disobedient at home”, “Has difficulty in settling to anything for more than a few moments”, “Is upset by new situation, by things happening for first time”, “Is bullied by other children”.

9
unity where this assumption is made is included. To ease interpretation, the Rutter indexes at each age are then reversed in sign (and so are increasing in mental health), and standardised to have a mean of zero and a standard deviation of one.

As an additional measure of mental health, we use data from the Bristol Social Adjustment Guide (BSAG), a measure of social skills used widely in the economics literature (Carneiro et al., 2007). The guide consists of a large number of phrases which describe a child’s behaviour, often increasing in severity (Stott, 1958). When cohort members were aged 7 and 11, the teachers of cohort members completed the guide by underlining the items of behaviour they thought related to the child. The items can be grouped into 12 domains, and maladjustment scores can be derived for each of these domains by summing the number of items underlined in each domain. For our measure of mental health, we use a total measure of behavioural deviance constructed as a sum of the scores of all domains. Again to ease interpretation, the BSAG total scores were reversed in sign, and were standardised to have a mean of zero and a standard deviation of one.

The distributions of each of the mental health measures at age 7 and 11 are show in Figure 2.

5.1.3. PREMATURE MORTALITY

Our second outcome of interest is premature mortality. Deaths of cohort members are identified regularly by those responsible for the running of the NCDS. Deaths are identified via multiple sources including death certificates from the National Health Service Central Register (NHSCR), NCDS death cards, and relatives and friends of cohort members who are contacted via telephone, letter and email during survey activities and cohort maintenance work. Stillbirths and neonatal deaths were identified by NCDS interviewers who remained in close contact with the families in the early post-natal period (Johnson and Brown, 2015a). Data on the exact month and year of death were obtained through augmenting the publically-available NCDS dataset with secure access data under special licence (Institute of Education, 2015a). The identification of deaths is unaffected by attrition from the NCDS, meaning that this dataset includes all deaths occurring prior to 31 December 2013. In only 6 cases is the date of death unknown (Johnson and Brown, 2015b). To examine the longevity effects of in-utero exposure to influenza, we construct a range of dummy variables relating to stillbirth/neonatal death, death within 28 days of birth, and death prior to each age of follow-up: 7, 11, 16, 23, 33, 42, 46, 50, and 55.

5.1.4. CHILD-LEVEL CONTROLS

We include a range of variables included in Kelly (2011): social class and schooling of both parents, maternal age and its squared value, maternal height, the number of persons per room, and a measure of the degree of maternal smoking prior to pregnancy.

However, Kelly (2011) only controls for characteristics that “plausibly influence the probability of infection” and therefore do not control for child characteristics such as gender and ethnicity, which have been found to be determinants of a range of life-course outcomes. However, the identification strategy relies on geographical variation in exposure and not solely on individual-level exposure, and so any differences (even by chance) in the gender and ethnic composition of local authorities are likely to result in biased results. We control for gender and ethnicity, as well as a range of additional pre-determined variables which we think are likely to be predictive of our outcomes. These include father’s age and its squared value, the number of older siblings, maternal marital status, and a categorical variable based on BMI which classifies mothers as being of a normal weight, underweight, over-
overweight, or obese. We also control for the mothers experience in previous pregnancies, including separate dummies for whether the mother had experienced a previous abortion, premature birth, large birth, stillbirth and a complication in pregnancy. As a measure of health investment we also control for whether the mother visited a neo-natal unit prior to 7 weeks into pregnancy. This visit is exogenous given that it pre-dates the epidemic and is highly correlated with total number of visits.

A more detailed description of controls can be found in Table 1.

5.4. LOCAL AUTHORITY INFLUENZA EXPOSURE

Throughout the 1950’s and 60’s, law did not require physicians to collect data on influenza cases, and so local authority data on influenza infection rates are not available. Instead, we follow Kelly (2011) in using local authority pneumonia infection rates. These serve as a good proxy for rates of influenza as, given they are clinically related, influenza can lead to the onset of pneumonia. In addition, trends in pneumonia infections closely follow trends in influenza deaths around the epidemic period (Hunter and Young, 1971).

Weekly data on pneumonia notifications by local authority were obtained from the Registrar Generals of England and Wales, and of Scotland (Registrar General for England and Wales, 1957; Registrar General for Scotland, 1957). To proxy influenza rates in the epidemic period, we construct a ratio of the total number of pneumonia notifications between September and November 1957 to the population size of the local authority. Infection rates in the pre-epidemic period are proxied using the same ratio but with the number of pneumonia notifications taken as the average notifications in the same months of 1956 and 1955. Consistent with other reports (Jackson, 2009), the deviations in pneumonia notifications from their pre-epidemic levels are distributed normally, departing from the long-run trend in September, and returning to trend at the end of November, with a peak in mid-October.

5.5. LOCAL AUTHORITY CHARACTERISTICS

Kelly (2011) control for a range of local authority characteristics derived from 1956 General’s Returns and the 1951 Census. These include, (i) the stillbirth rate, (ii) the proportion of households living with >1 person per room, (iii) the proportion of the male working population in an unskilled occupation, (iv) the proportion of men leaving school aged 16 or older, (v) population density, and (vi) the proportion of the population aged 65+. Given a current lack of access to this data, we cannot include these measures directly. However, data on (ii), (iii), and (iv) is available at the individual level in the NCDS dataset. Under the assumption that the NCDS cohort members in each local authority are a random sample of each local authority population, we use the aggregated individual-level variables from the NCDS as proxies for local authority proportions. We additionally use a count of NCDS cohort members per local authority as a proxy for population size.

Local authority level variables were matched to NCDS cohort members via identifiers of their local authority of birth, which were obtained under special licence (Institute of Education, 2015b).

5.6. DERIVING THE FINAL ESITEMATION SAMPLES

The process describing how the estimation samples are derived for each outcome variable is depicted in Figure 3. Throughout the life of the NCDS, 18,558 individuals completed at least one interview. We began by dropping individuals with no PMS record, studies suggest that the ratio of pneumonia notifications to influenza cases is approximately one to 417 (Hunter and Young, 1971).

Special thanks to Elaine Kelly for the provision of this data.
reducing the sample to 17,417 cohort members. We then dropped observations where data was missing on the local authority at birth or where data was missing on local authority influenza rates, resulting in the loss of a further 26 observations. We then excluded cases where data on individual-level exposure to influenza was missing. In its update of the PMS variables, the CLS failed to convert data for 446 cohort members. 426 of these relate to multiple births (212 twin-pairs and 4 sets of triplets), whose data where dropped in line with National Birthday Trust Fund policy. Dropping these observations reduced the sample size further to 16,944. When estimating the effect of in-utero exposure to influenza on mental health outcomes at ages 7 and 11, we dropped individuals who failed to complete the age 7 and age 11 surveys, respectively. Estimating the degree to which childhood mental health captures the effect of in-utero exposure to influenza on longevity outcomes requires that mental health at both age 7 and 11 is available, and so when estimating these effects, we dropped individuals who failed to complete both the age 7 and age 11 questionnaires. For all outcomes, we dropped individuals with missing data on the outcomes. We finally dropped observations with missing data on any of the control variables.

Implementing this process resulted in final estimation samples of 12,004, 12,333, 11,171, 11,416, 14,628, and 10,198 when estimating effects on the Rutter index at age 7, the BSAG total score at age 7, the Rutter index at age 11, the BSAG total score at age 11, longevity at all ages, and the mediating role of mental health, respectively.

6. RESULTS

6.1. SUMMARY STATISTICS

The summary statistics for the control variables, stratified by whether a cohort member experienced influenza in-utero, can be found in Table 1. A substantial proportion of cohort members were exposed to influenza in-utero, with 13.32% of NCDS mothers reporting that they experienced influenza during pregnancy. There is remarkably little difference in average values of child and family background characteristics between those who experienced fetal exposure to influenza and those who did not. However, those who experienced the fetal health shock were more likely to have mothers who had previous experience of a premature birth, a large birth, and a previous pregnancy-related complication. They were also less likely to have a father in the highest two social classes, and less likely to have mothers who were a normal/healthy weight prior to pregnancy.

In terms of the local authority characteristics, those who experienced in-utero exposure to influenza tended to be born in local authorities with a smaller number of NCDS cohort members, suggesting they were more likely to be born into areas of lower population size. They were also more likely to be born in local authorities with less cramped living conditions, with them being less likely to live in areas with a large proportion of individuals living in households with >1 person per room. Surprisingly, cohort members whose mothers contracted influenza during pregnancy were born into local authorities with lower average pneumonia rates both in the pre-epidemic and epidemic periods, with similarly small differences present in the pneumonia rate increases between these periods. This is concerning given that our identification strategy relies on the geographical variations in the degree of deviations in infection rates from pre-epidemic levels differing by individual influenza infection status.

Summary statistics for the outcomes suggest that being exposed to influenza in-utero had very little effect on mental health in childhood (Table 2), although at both ages 7 and 11 and for both measures, mean scores for those who were exposed to influenza in-utero were lower than that of those who weren’t exposed. The impact on the longevity of the cohort seems much

---

9 For ease of presentation, we present these only for those in the estimation sample for the mortality outcomes.
stronger, although the differential in cumulative deaths between the two groups (see Figure 4) seems to be driven by strong positive effects on the probability of death at birth.

6.2. REGRESSION RESULTS

6.1.1. CHILDHOOD MENTAL HEALTH

Table 3 shows the effects of in-utero exposure to influenza on mental health outcomes at age 7 and 11, estimated from our baseline specification, where $$f_{\text{flu}}$$ is assumed to be exogenous. For each measure at each age, the first column presents the unconditional effect, the second column presents effects after controlling for local authority characteristics and region fixed effects, and the third column contains the effects after additionally controlling for child-level characteristics. As outcomes are standardised, the coefficients can be interpreted on the standard deviation scale. At age 7, irrespective of the measure used and the controls included, the epidemic failed to significantly reduce mental health. At age 11, when only local authority characteristics and region fixed effects were controlled for, fetal influenza exposure was associated with a 0.06 standard deviation reduction in mental health as measured by the Rutter index, which was statistically significant at the 5% level. However, this effect became statistically insignificant once child characteristics were controlled for.

Table 4 reports estimates from the IV approach which exploits the plausibly exogenous variations in epidemic-induced deviations in influenza rates from their long-run trends across local authorities. The effects found here are puzzling. At age 7, estimates of the effect of in-utero exposure to influenza vary widely dependent on the measure used, with results suggesting large negative effects of influenza on the Rutter index, but large positive effects on the BSAG total score. However, due to considerably large standard errors, these large magnitudes do not translate into statistical significance. The opposite result was found at age 11, with IV estimates suggesting large positive, but statistically insignificant, effects on mental health on the Rutter index and large negative effects on mental health on the BSAG total score. Diagnostic tests for the first stage regressions indicate the reason for this. The F-statistic indicating the strength of the instruments is highly insignificant in all specifications, indicating that a higher intensity of the epidemic in cohort members’ local authority of birth does not increase the probability of in-utero exposure to influenza. In addition, results from the Durbin-Wu tests indicate that there is no endogeneity in the second stage regression, suggesting that individual-level influenza exposure is exogenous in our baseline specification. Again this is puzzling given a substantial body of literature suggesting multiple sources of endogeneity. However, it must be noted that weak instruments reduce the accuracy of these endogeneity tests (Tchatoka, 2012).

6.1.2. LONGEVITY

The effects of fetal exposure to influenza on the probability of death by multiple ages are presented in Tables 5, 6 and 7. Results from the baseline specification with all controls included (Table 5) indicate that in-utero exposure to influenza significantly reduces age of death. A strong effect is found at birth, where fetal exposure to influenza increases the probability of being stillborn or dying in a neonatal unit by just over one percentage point, and this effect is statistically significant at a 1% level. The magnitude of this effect, although fluctuating, remains around one percentage point across the life-course, and maintains statistical significance at 28 days and at ages 7, 11, 23, 33, 42, and 46. Given this constant magnitude, it seems likely that the significant results for ages later in the life-course are driven mainly by the differential at birth. To investigate this, we estimate the same specification but control for death at birth (Table 7). Results confirm this hypothesis, with estimates of the effect of
exposure on the probability of death at all ages beyond birth becoming statistical insignificant and close to zero in magnitude once death at birth is included.

As with the child mental health outcomes, IV estimates of the longevity effects of the in-utero exposure to influenza are strange (Table 6). The coefficients are very large in magnitude and, as we are dealing with binary outcomes, also implausible. The models predict that exposure to influenza in-utero reduces the probability of death by over 100% at some ages. However, these are again insignificant, driven by the weak instrument problem (the p-value of the first stage F-statistic exceeds 0.5). Again, in the majority of specifications, the Durbin-Wu test statistic is statistically insignificant, suggesting the exogeniety of individual-level in-utero exposure to influenza in the baseline specification.

6.1.3. MENTAL HEALTH AND LONGEVITY

Despite finding limited effects on mortality beyond childbirth, we study the degree to which mental health can explain these longevity effects (Table 8). Repeating our basic mortality regressions only on the sample of individuals with data on mental health outcomes at age 7 and 11 does not change our findings substantially. The inclusion of the mental health variables does little to change these coefficients, indicating that mental health is not an important pathway through which fetal health conditions affect mortality.

7. DISCUSSION

This study represents one of the first attempts to investigate the causal effects of fetal exposure to infection on child mental health outcomes. We use plausibly exogenous geographical variation in the intensity of the 1957 Asian influenza epidemic to identify these effects in a cohort of British children. We also add to the methodological literature in this area by making use of individual-level exposure to infection. However, our preferred specification is one where in-utero exposure to influenza is treated as exogenous, largely driven by the poor performance of local authority-level epidemic intensity as an instrument for individual-level exposure.

Using this preferred specification, we find that in-utero exposure to influenza had no impact on mental health at either age 7 or age 11. This is surprising given that an effect on cognitive development has been found in the same cohort (Kelly, 2011), and given the variety of pathways through which fetal health conditions can affect mental health (Schlotz and Phillips, 2009). However, our findings are consistent with one of the only other studies to examine child mental health outcomes in the fetal origins literature. Maselko et al. (2015) examine the effects of the Thinking Healthy Programme, a cognitive behavioural therapy intervention for pregnant mothers with prenatal depression. Despite the trial generating a significant reduction in the probability of mothers experiencing depression a year after the intervention, they find that this exogenous change in maternal depression had little impact on their offspring’s socio-emotional development up to age 7. However, our lack of findings may be due to the crudeness of our measures of mental health. Mental health is multi-faceted, and so our null finding may mask significant effects in particular aspects of mental well-being. One option for further work would be to follow previous literature in using principle components analysis to split the Rutter index into separate externalising and internalising behaviour scales (Blanden et al., 2007). A similar approach has also been used for the BSAG (Jokela et al., 2009). We could also study effects on the BSAG syndrome scores, although some have cautioned against this (Shepherd, 2013).

We also examine the effects of in-utero exposure to influenza on the longevity of the same cohort. Our findings suggest a substantial impact of fetal exposure on the probability of being stillborn. This is consistent with the medical literature that states...
that influenza can lead to complications in late pregnancy such as septic shock, meningitis and encephalitis which can subsequently lead to stillbirths. Currently, all women in the UK are advised to have a flu vaccine to protect against these risks (NHS, 2015). Our results highlight the importance of this. We also find that the magnitude of the effect remained relatively constant up to age 55, indicating a persistence in the survival advantage for those not exposed to influenza in-utero, but no additional survival advantage for those living beyond infancy. This is consistent with the fetal origins hypothesis which states that the health effects of fetal adversity may remain latent until old age, when the onset of physical disease becomes prevalent. Indeed, age 55 may not be sufficiently old for us to identify these effects. This is consistent with an evaluation of the long-term health effects of a neonatal program in Sweden, where the mortality advantage for the treated individuals began in infancy and remained constant throughout adulthood, before increasing again at age 70 (Bhalotra et al., 2015a). It would be interesting to see if similar results are found in the NCDS cohort when data from later follow-ups become available.

A further interesting implication of this study is the failure of geographical fluctuations in epidemic intensity to predict individual-level exposure. Our results suggest that the two are at best weakly correlated. This is concerning given that many of the empirical tests of the fetal origins hypothesis use geographical variation in pandemic severity (or other shocks) to identify the effects the in-utero environment on later-life outcomes. The assumption that area-level exposure predicts individual-level exposure is therefore key. Later studies should examine this link more carefully.

However, this latter recommendation is weakened by potential shortfalls in our own empirical strategy. Due to the absence of data on area-level influenza cases during and before the epidemic period, pneumonia rates were instead used as a proxy, under the assumption that there was both a medical and statistical link between their prevalences. A failure of this assumption may be the reason for the weakness of the instrument. Additionally, the individual-level data on influenza cases in the NCDS mothers is only a new release and so may be subject to errors. Indeed, the rate of influenza exposure in the NCDS cohort is just over 13%, well below the rates of approximately 30% suggested in reports published around the time of the epidemic (Woodall et al., 1958). If this represents under-reporting, and if this under-reporting varies systematically across local authorities, this may weaken the link between individual-level infection and area-level exposure rates. Furthermore, this link may be weakened by non-random selection into the NCDS sample. Inclusion in all estimation samples required mothers of NCDS cohort members to complete the NCDS birth survey. It is possible that the contraction of influenza could have reduced the probability of survey completion, reducing completion rates in areas of high epidemic intensity. To investigate this, future work will first assess whether there are differences between local authority-level birth rates in the population and those implied by the numbers of NCDS cohort members in each local authority, before testing whether these differences are correlated with epidemic intensity. A significant correlation would suggest non-random selection.

In addition, one of the mechanisms outlined in section 3 could invalidate the use of area-level pandemic intensity as an instrument for individual-level exposure to influenza in-utero. It was suggested that in-utero exposure to influenza could affect childhood mental health through a deterioration in economic conditions in areas of high pandemic intensity. However, both those with mothers contracting and those with mothers not contracting influenza could be affected by these adverse economic conditions. As a result, this mechanism suggests a direct effect of area-level pandemic intensity on mental health outcomes, independent of its effects on individual-level exposure, violating the excludability assumption. This again may raise a problem in the wider fetal origins literature, where area-level exposure, may this be to a pandemic or to some other shock, is implicitly used as an ‘instrument’ for individual-level fetal adversity in many identification strategies. This potential problem may mean that
many studies in this area are not testing the fetal origins hypothesis, but instead testing the overall effect of a pandemic (or other shock). Later studies should examine the implications this problem may have for their identification strategies.

Furthermore, our results may be subject to several other forms of bias. First is bias due to unobserved local authority-level characteristics. Although the inclusion of region fixed effects restricts this to bias caused by characteristics which vary both over time and within region, unobservables likely remain. One such unobservable could be the local quality of healthcare services, which could be correlated with both local-level infection rates and the health outcomes of cohort members. Future work will strive to include proxy measures of this, such as the local percentage of deaths from unknown causes used in Neelsen and Stratmann (2012), as well as the other local authority characteristics included in Kelly (2011). Secondly, we fail to control for maternal mental health problems. This may be correlated with both the probability of contracting (or at least reporting) influenza during pregnancy. Mental health problems also have a strong hereditary component and thus maternal mental health problems are also likely to be correlated with childhood mental health problems (Johnston et al., 2013). Further work will make use of a question in the PMS asking mothers whether they were treated for a psychiatric disorder during pregnancy, although there is a concern that this may itself be an outcome on contracting influenza, and would therefore constitute a bad control. Thirdly, there are concerns regarding the quality of the data on deaths. Although it is claimed that the NCDS has recorded all but six deaths in the NCDS cohort, the data suggests that rates of death are much lower for those who emigrate from Great Britain. Further work will investigate whether this is explained by a healthy migrant effect or whether this is a selection issue, which will require further correspondence with the Centre for Longitudinal Studies. Fourthly, there is a concern that the negative health effects may be inflated due to internal migration in response to the epidemic. It may be that unobservably better mothers, which invest more in the health of their offspring, move away from areas of high epidemic intensity, thus reducing the probability of influenza infection. However, evidence suggests that the spread of the disease was quick, with general practitioners being “overawed at the suddenness of its outset” (Jackson, 2009). This, along with the randomness in the spread of the epidemic (Hunter and Young, 1971), gives us confidence that it would be difficult to avoid infection through migration, and therefore that internal migration will have little effects on our results. Finally, our null findings for mental health could be driven by mortality selection. It is suggested in the medical literature that it is the already unhealthy infants who succumb to infection. As a result, the mortality differentials prior to age 7 could have pruned the unhealthy children from the population of individuals exposed to influenza, increasing average levels of mental health in the infected group relative to those in the non-infected group. This is likely to make our estimates a conservative estimate of the true effect of in-utero exposure to influenza on mental health in childhood.

Despite these limitations, we believe this study, upon refinement, makes a substantial contribution to the fetal origins literature, both in terms of investigations into effects on previously under-explored outcomes, and in the use of new methodology. Future work will build further on the literature, by using data from the offspring of the NCDS cohort members to study the intergenerational effects in-utero exposure to influenza. In addition, subgroup analysis and quantile regressions will also be performed to identify potential heterogeneity in the effects of in-utero exposure to influenza across different subsets of the population.
8. BIBLIOGRAPHY


Black, S., Bütkofer, A., Devereux, P., Salvanes, K., 2013. This is only a test? long-run impacts of prenatal exposure to radioactive fallout.


Figure 1: Maternal behavior and health conditions, the fetal environment, and offspring mental health

Figure 2: Distributions of the child mental health outcomes at age 7 and 11
Figure 3: Derivation of the estimation sample for each outcome variable

Figure 4: Cumulative deaths from birth to age 55, by influenza status
### Table 1: Descriptive statistics of control variables, by influenza status

<table>
<thead>
<tr>
<th>Variable</th>
<th>With influenza (13.32%)</th>
<th>Without flu (86.68%)</th>
<th>Variable</th>
<th>With influenza (13.32%)</th>
<th>Without flu (86.68%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child-level characteristics</strong></td>
<td></td>
<td></td>
<td>Persons per room</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother schooled past leaving age</td>
<td>0.252 .</td>
<td>0.257 .</td>
<td>&lt;1</td>
<td>0.691 .</td>
<td>0.687 .</td>
</tr>
<tr>
<td>Father schooled past leaving age</td>
<td>0.583 .</td>
<td>0.594 .</td>
<td>1-1.5</td>
<td>0.174 .</td>
<td>0.171 .</td>
</tr>
<tr>
<td>Mother’s age at birth</td>
<td>27.241 5.717</td>
<td>27.556 5.616</td>
<td>1.5-2</td>
<td>0.093 .</td>
<td>0.094 .</td>
</tr>
<tr>
<td>Mother’s height (inches)</td>
<td>63.435 2.530</td>
<td>63.403 2.498</td>
<td>&gt;2</td>
<td>0.042 .</td>
<td>0.048 .</td>
</tr>
<tr>
<td>Male</td>
<td>0.511 .</td>
<td>0.518 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non white</td>
<td>0.008 .</td>
<td>0.011 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of older siblings</td>
<td>1.336 1.511</td>
<td>1.221 1.458</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother married</td>
<td>0.992 .</td>
<td>0.991 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous abortion</td>
<td>0.124 .</td>
<td>0.121 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous premature birth (&lt;5lb,9oz)</td>
<td>0.095 .</td>
<td>0.081 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous large birth (&gt;8lb,14oz)</td>
<td>0.106 .</td>
<td>0.091 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous stillbirth/neonatal death</td>
<td>0.051 .</td>
<td>0.054 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complication in previous pregnancy</td>
<td>0.093 .</td>
<td>0.086 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antenatal visits within 7 weeks</td>
<td>0.043 .</td>
<td>0.041 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mother’s occupational social class</strong></td>
<td></td>
<td></td>
<td>Weight prior to pregnancy (based on BMI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.613 .</td>
<td>0.627 .</td>
<td>Normal/healthy</td>
<td>0.720 .</td>
<td>0.731 .</td>
</tr>
<tr>
<td>I or II: Professional, managerial or technical</td>
<td>0.047 .</td>
<td>0.046 .</td>
<td>Underweight</td>
<td>0.027 .</td>
<td>0.039 .</td>
</tr>
<tr>
<td>III: Skilled manual/non-manual</td>
<td>0.214 .</td>
<td>0.211 .</td>
<td>Overweight</td>
<td>0.204 .</td>
<td>0.189 .</td>
</tr>
<tr>
<td>IV: Partly skilled</td>
<td>0.101 .</td>
<td>0.092 .</td>
<td>Obese</td>
<td>0.048 .</td>
<td>0.042 .</td>
</tr>
<tr>
<td>V: Unskilled</td>
<td>0.024 .</td>
<td>0.025 .</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Father’s occupational social class</strong></td>
<td></td>
<td></td>
<td>Local authority variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I: Professional</td>
<td>0.040 .</td>
<td>0.047 .</td>
<td>Pneumonia notifications per1,000 - Epidemic period</td>
<td>0.420 .</td>
<td>0.420 .</td>
</tr>
<tr>
<td>II: Managerial or technical</td>
<td>0.104 .</td>
<td>0.133 .</td>
<td>Pneumonia notifications per1,000 - Pre-epidemic period (1955/56)</td>
<td>0.119 .</td>
<td>0.155 .</td>
</tr>
<tr>
<td>III: Skilled manual/non-manual</td>
<td>0.626 .</td>
<td>0.591 .</td>
<td>Number of NCDS cohort members</td>
<td>362.859</td>
<td>460.695</td>
</tr>
<tr>
<td>IV: Partly skilled</td>
<td>0.119 .</td>
<td>0.117 .</td>
<td>Proportion living in household with &gt;1 persons per room</td>
<td>0.315 .</td>
<td>0.322 .</td>
</tr>
<tr>
<td>V: Unskilled</td>
<td>0.092 .</td>
<td>0.093 .</td>
<td>Proportion of men leaving school aged 16+</td>
<td>0.097 .</td>
<td>0.098 .</td>
</tr>
<tr>
<td>Unclassifiable/no father</td>
<td>0.020 .</td>
<td>0.019 .</td>
<td>Proportion of male working population unskilled</td>
<td>0.230 .</td>
<td>0.231 .</td>
</tr>
</tbody>
</table>

*All figures to 3 d.p.*
<table>
<thead>
<tr>
<th>Variable</th>
<th>With influenza (13.32%)</th>
<th>Without flu (86.68%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Childhood mental health</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rutter behaviour index: age 7</td>
<td>-0.033</td>
<td>1.024</td>
</tr>
<tr>
<td>BSAG total score: age 7</td>
<td>0.005</td>
<td>0.978</td>
</tr>
<tr>
<td>Rutter behaviour index: age 11</td>
<td>-0.047</td>
<td>1.010</td>
</tr>
<tr>
<td>BSAG total score: age 11</td>
<td>-0.009</td>
<td>1.009</td>
</tr>
<tr>
<td><strong>Longevity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stillbirth</td>
<td>0.028</td>
<td>.</td>
</tr>
<tr>
<td>Dead within 28 days</td>
<td>0.030</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 7</td>
<td>0.039</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 11</td>
<td>0.040</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 16</td>
<td>0.042</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 23</td>
<td>0.046</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 33</td>
<td>0.050</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 42</td>
<td>0.058</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 46</td>
<td>0.065</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 50</td>
<td>0.073</td>
<td>.</td>
</tr>
<tr>
<td>Dead by age 55</td>
<td>0.078</td>
<td>.</td>
</tr>
</tbody>
</table>

*All figures to 3 d.p.*
### Table 3: Regression results from the basic specification: Childhood mental health

<table>
<thead>
<tr>
<th></th>
<th>Rutter behaviour index</th>
<th>BSAG total score</th>
<th>Rutter behaviour index</th>
<th>BSAG total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetal influenza exposure</td>
<td>-0.0369</td>
<td>-0.0402</td>
<td>-0.0330</td>
<td>-0.0213</td>
</tr>
<tr>
<td></td>
<td>(-1.34)</td>
<td>(-1.46)</td>
<td>(-1.22)</td>
<td>(-0.82)</td>
</tr>
<tr>
<td>LA controls and region FE</td>
<td>No Yes Yes</td>
<td>No Yes Yes</td>
<td>No Yes Yes</td>
<td>No Yes Yes</td>
</tr>
<tr>
<td>Child-level controls</td>
<td>No No Yes Yes</td>
<td>No No Yes Yes</td>
<td>No No No Yes</td>
<td>No No No Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>11980 11980 11980</td>
<td>12209 12209 12209</td>
<td>11148 11148 11148</td>
<td>11394 11394 11394</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.000 0.008 0.047</td>
<td>0.000 0.002 0.064</td>
<td>0.000 0.006 0.041</td>
<td>0.000 0.003 0.084</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parenthesis

### Table 4: Regression results from the instrumental variable approach: Childhood mental health

<table>
<thead>
<tr>
<th></th>
<th>Rutter behaviour index</th>
<th>BSAG total score</th>
<th>Rutter behaviour index</th>
<th>BSAG total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetal influenza exposure</td>
<td>-8.737</td>
<td>-4.381</td>
<td>3.422</td>
<td>12.44</td>
</tr>
<tr>
<td></td>
<td>(-0.76)</td>
<td>(-0.45)</td>
<td>(0.47)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>LA controls and region FE</td>
<td>No Yes Yes</td>
<td>No Yes Yes</td>
<td>No Yes Yes</td>
<td>No Yes Yes</td>
</tr>
<tr>
<td>Child-level controls</td>
<td>No No Yes Yes</td>
<td>No No No Yes</td>
<td>No No No Yes</td>
<td>No No No Yes</td>
</tr>
<tr>
<td>F-statistic p-value</td>
<td>0.449</td>
<td>0.636</td>
<td>0.415</td>
<td>0.713</td>
</tr>
<tr>
<td>Durbin-Wu p-value</td>
<td>0.042</td>
<td>0.579</td>
<td>0.383</td>
<td>0.240</td>
</tr>
<tr>
<td>Observations</td>
<td>11980 11980 11980</td>
<td>12209 12209 12209</td>
<td>11148 11148 11148</td>
<td>11394 11394 11394</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parenthesis

### Table 5: Regression results from the basic specification: longevity outcomes

<table>
<thead>
<tr>
<th></th>
<th>Stillbirth</th>
<th>28 days</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Age23</th>
<th>Age 33</th>
<th>Age 42</th>
<th>Age 46</th>
<th>Age 50</th>
<th>Age 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetal influenza exposure</td>
<td>0.0108**</td>
<td>0.0122**</td>
<td>0.00979*</td>
<td>0.00984*</td>
<td>0.00883</td>
<td>0.0112*</td>
<td>0.0134*</td>
<td>0.0140*</td>
<td>0.0129*</td>
<td>0.0111</td>
<td>0.0108</td>
</tr>
<tr>
<td></td>
<td>(2.87)</td>
<td>(2.74)</td>
<td>(2.08)</td>
<td>(2.06)</td>
<td>(1.82)</td>
<td>(2.16)</td>
<td>(2.46)</td>
<td>(2.40)</td>
<td>(2.12)</td>
<td>(1.74)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>LA controls and region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Child-level controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.086</td>
<td>0.143</td>
<td>0.176</td>
<td>0.173</td>
<td>0.168</td>
<td>0.152</td>
<td>0.145</td>
<td>0.126</td>
<td>0.113</td>
<td>0.104</td>
<td>0.098</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parenthesis
### Table 6: Regression results from the instrumental variable approach: longevity outcomes

<table>
<thead>
<tr>
<th></th>
<th>Stillbirth</th>
<th>28 days</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Age23</th>
<th>Age 33</th>
<th>Age 42</th>
<th>Age 46</th>
<th>Age 50</th>
<th>Age 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetal influenza exposure</td>
<td>-1.035</td>
<td>-2.600</td>
<td>-1.158</td>
<td>-1.009</td>
<td>-0.517</td>
<td>-0.847</td>
<td>-1.095</td>
<td>-1.636</td>
<td>-1.965</td>
<td>-1.787</td>
<td>-1.101</td>
</tr>
<tr>
<td></td>
<td>(-0.29)</td>
<td>(-0.28)</td>
<td>(-0.26)</td>
<td>(-0.26)</td>
<td>(-0.21)</td>
<td>(-0.25)</td>
<td>(-0.28)</td>
<td>(-0.30)</td>
<td>(-0.30)</td>
<td>(-0.31)</td>
<td>(-0.30)</td>
</tr>
<tr>
<td>LA controls and region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Child-level controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-statistic p-value</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
<td>0.7772</td>
</tr>
<tr>
<td>Durbin-Wu p-value</td>
<td>0.4163</td>
<td>0.039</td>
<td>0.371</td>
<td>0.4348</td>
<td>0.6381</td>
<td>0.5259</td>
<td>0.4067</td>
<td>0.3201</td>
<td>0.2798</td>
<td>0.3972</td>
<td>0.5957</td>
</tr>
<tr>
<td>Observations</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parenthesis

### Table 7: Regression results from the basic specification: longevity outcomes, controlling for whether stillborn

<table>
<thead>
<tr>
<th></th>
<th>Stillbirth</th>
<th>28 days</th>
<th>Age 7</th>
<th>Age 11</th>
<th>Age 16</th>
<th>Age23</th>
<th>Age 33</th>
<th>Age 42</th>
<th>Age 46</th>
<th>Age 50</th>
<th>Age 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fetal influenza exposure</td>
<td>0.0108**</td>
<td>0.00228</td>
<td>0.000367</td>
<td>0.000439</td>
<td>-0.000548</td>
<td>0.00181</td>
<td>0.00411</td>
<td>0.00480</td>
<td>0.00386</td>
<td>0.00206</td>
<td>0.00181</td>
</tr>
<tr>
<td></td>
<td>(2.87)</td>
<td>(0.80)</td>
<td>(0.11)</td>
<td>(0.13)</td>
<td>(-0.15)</td>
<td>(0.45)</td>
<td>(0.94)</td>
<td>(0.99)</td>
<td>(0.74)</td>
<td>(0.37)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>Stillborn</td>
<td>0.924***</td>
<td>0.875***</td>
<td>0.873***</td>
<td>0.871***</td>
<td>0.867***</td>
<td>0.858***</td>
<td>0.850***</td>
<td>0.844***</td>
<td>0.837***</td>
<td>0.832***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(172.99)</td>
<td>(118.84)</td>
<td>(118.14)</td>
<td>(116.91)</td>
<td>(115.19)</td>
<td>(111.47)</td>
<td>(108.07)</td>
<td>(105.34)</td>
<td>(102.27)</td>
<td>(100.40)</td>
<td></td>
</tr>
<tr>
<td>LA controls and region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Child-level controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
<td>14603</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.086</td>
<td>0.606</td>
<td>0.510</td>
<td>0.498</td>
<td>0.478</td>
<td>0.435</td>
<td>0.397</td>
<td>0.343</td>
<td>0.305</td>
<td>0.275</td>
<td>0.258</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parenthesis
Table 8: Examining the degree to which childhood mental health captures the effect of in-utero exposure to influenza on longevity outcomes

<table>
<thead>
<tr>
<th>Age 16</th>
<th>Age 23</th>
<th>Age 33</th>
<th>Age 42</th>
<th>Age 46</th>
<th>Age 50</th>
<th>Age 55</th>
</tr>
</thead>
<tbody>
<tr>
<td>No MH</td>
<td>With MH</td>
<td>No MH</td>
<td>With MH</td>
<td>No MH</td>
<td>With MH</td>
<td>No MH</td>
</tr>
<tr>
<td>Fetal influenza exposure</td>
<td>-0.000474</td>
<td>-0.000515</td>
<td>0.000196</td>
<td>0.00184</td>
<td>0.00249</td>
<td>0.000233</td>
</tr>
<tr>
<td>Rutter index: Age 7</td>
<td>-0.000608</td>
<td>-0.000383</td>
<td>-0.00221</td>
<td>-0.00217</td>
<td>-0.00306</td>
<td>-0.00425</td>
</tr>
<tr>
<td>Rutter index: Age 11</td>
<td>-0.000333</td>
<td>-0.00221*</td>
<td>-0.00130</td>
<td>-0.00143</td>
<td>-0.00221</td>
<td>-0.00272</td>
</tr>
<tr>
<td>LA controls and region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Child-level controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>10175</td>
<td>10175</td>
<td>10175</td>
<td>10175</td>
<td>10175</td>
<td>10175</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.006</td>
<td>0.007</td>
<td>0.004</td>
<td>0.005</td>
<td>0.006</td>
<td>0.007</td>
</tr>
</tbody>
</table>

* p<0.05, ** p<0.01, *** p<0.001; t-statistics in parenthesis