Childhood Health & Education:
The Impact of Influenza at Different Ages on Education

Nolan J.C.L. Noble
University of Notre Dame

19 August 2013

Abstract:

This paper explores the latent effect of childhood health on education using influenza, a severe annual health threat to children. Elderly influenza deaths proxy children's exposure to the flu; children experience most of the morbidity, while the elderly are most likely to experience mortality. These two series are highly correlated. Overall influenza reduces schooling. However, cross-type immunity acquisition in children too young to attend school serves to increase their educational attainment. These gains are outweighed by losses suffered from influenza later in childhood. Individuals are 0.32%–1.85% less likely to report higher education when exposed to median influenza throughout childhood. This implies lost wages of $9.4–$37.4 billion.

JEL Classification: I10, 120

Keywords: Influenza, childhood health, immunity acquisition, educational attainment, timing

Acknowledgments: I thank Bill Evans, Molly Lipscomb, Kasey Buckles, and participants at the Midwest Economic Association, Western Economic Association International, and AcademyHealth annual meetings. This work was supported by NSF IGERT grant award #0504495 to the GLOBES graduate training program and by the Kellogg Institute for International Studies both at the University of Notre Dame. All remaining errors are my own.

1. Introduction

There is a growing body of work which suggests that health during childhood matters for educational attainment, with health insults reducing schooling.[1] Studies have demonstrated the basic relationship in both a developed county context (Almond, 2006, Bleakley, 2007, Bleakley, 2009, Black et al., 2007, Case et al., 2004, Currie, Neidell, 2005, Currie et al., 2009, Currie et al., 2010, Oreopoulos et al., 2008) and in emerging economies (Baird et al., 2011, Barreca, 2010, Burlando, 2009, Maccini, Yang, 2009, Miguel, Kremer, 2004). A good portion of this empirical literature has focused exclusively on health risks around infancy or in utero. Few studies have attempted to identify how the timing of health shocks throughout childhood affects education. An early example is Case et al. (2004) who examined the presence of chronic conditions at two points during childhood. Currie et al. (2010) is a more recent example, they examined mental health and injuries during five-year intervals. Understanding the impact of health shocks during childhood is important for a variety of reasons. First, much of the focus has been on prenatal conditions because of data availability. Birth conditions are frequently recorded in data sets and the availability of birth dates allows researchers to examine environmental conditions at birth. Health shocks can and frequently occur at any time and hence understanding when children are most vulnerable is important. Second, understanding how the timing of shocks is important helps identify the mechanism through which health shocks during different periods of childhood affect education. For example, if there are large impacts of health shocks during school, then missed schooling is a likely mechanism. Instead, if the impact of health shocks are primarily in children younger than school age, then the pathway is likely through a reduction in the permanent stock of health.

In this paper I investigate the effects of negative health shocks throughout childhood on education. The childhood health shock I use is exposure to influenza. I begin by presenting the impact of exposure at each age during childhood. Since I have exposure rather than morbidity, the impact of exposure at each age provides a partial picture. To obtain a complete picture, I use the probabilities of morbidity at each age to yield an aggregate impact of childhood influenza on educational attainment. This paper contributes to the literature on childhood health and education by identifying differences across ages in the impact of health shocks, specifically examining health during the entirety of childhood, rather than focusing on *in utero* or infant

[1] See Almond & Currie (2011) or Currie (2011) for recent reviews.

health. I know of no other existing work that isolates the effect of health shocks in annual age cohorts of childhood.

Influenza is a severe childhood disease that lends itself to study of its impact on educational attainment. Children's experience with influenza is much worse than adults (Beigel, 2008, Monto, 2008, Simonsen, 2001, Wilschut et al., 2006). Incidence rates are much higher. They experience more symptoms, which are more severe and last longer (up to ten days). Pneumonia, hospitalization, and other complications are not uncommon. Being an infectious disease and considering influenza's impact throughout childhood allows for the exploration of an immunity acquisition channel. Like most infectious diseases, immunity to influenza is possible and can be acquired through exposure. Additionally, cross-(sub) type immunity is possible (Epstein, 2006, Glezen, Couch, 1997). Early exposure when the educational costs of exposure are not large can possibly provide a benefit later on. Consider two otherwise identical children: one exposed to influenza before school (subsequently obtaining immunity) and one not. These children are then re-exposed to influenza during school. The child lacking immunity will become sick and miss school, while the other remains healthy. This implies that the child who gained immunity benefited from the early exposure. This paper is able to capture the immunity benefits of early exposure to influenza because it observes the differential impact from the initial exposure to influenza on educational attainment across childhood.

To conduct this study I combine data measuring influenza exposure and the corresponding outcomes of the afflicted children. A natural way to think about this is what is a child's contraction of the flu on their outcomes. This is subject to omitted variables bias concerns because idiosyncrasies can enhance/detract the probability of illness. Additionally, this is a limited impact, as one kid's flu can have spillovers to others. One way to avoid these problems is to instead use exposure to influenza. Exposure is ideal to use in that it is free from these simultaneity concerns and it is policy relevant. It captures one's risk and the risk to all other kids. I use elderly influenza death rates constructed at the state-year level as a proxy for virulence. This measure captures influenza's considerable variation in virulence across time and space stemming from the virus's volatile nature. I document that child influenza morbidity is highly correlated with elderly influenza deaths. The influenza exposure data is then combined at the state-year level with individual-level data from the 2000 U.S. Census 5-Percent Public Use Micro Samples (PUMS).

In contrast to much of the existing literature, I find that flu exposure around infancy has no impact on educational attainment. Interestingly, flu exposure prior to entering school increases education, likely through an immunity acquisition channel. Upon entering school, influenza reduces education. The time that matters most are grades three through six and twelve, ranging from a reduction in the probability of graduating from high school by 0.21%–0.53% and for obtaining a bachelor's degree 0.86%–5.15% due to exposure in a single year. Overall, the negative effects of influenza overpower the benefits of immunity acquisition. I use the estimated probability of experiencing a shock from Lemaitre & Carrat (2010) to estimate the effect of influenza exposure during the entirety of childhood. Conditional on the level of influenza activity (low to epidemic) during childhood and the measure of educational attainment, total childhood influenza exposure decreases schooling by 0.13%–2.96%. This verifies that timing matters, specifically, influenza exposure during 3^{rd}–6^{th} and 12^{th} grades are the most important for education. The results highlight that influenza is having an economically meaningful impact on education. For comparison, Almond (2006) finds that a reduction in maternal nutrition due to the 1918 Spanish Flu pandemic reduced the probability of graduating from high school by two-percent.

The positive impact of influenza exposure prior to entering school has clear policy implications: low cost (in educational terms) immunity acquisition benefits education. It is possible that other infectious diseases may exhibit a similar beneficial (in terms of education) period of exposure and or morbidity, but that is beyond the scope of this paper. Generalizing these results beyond the sample studied imply that influenza reduced the combined number of high school and college graduates by 743,558 –2,974,232. These lost degrees result in lost wages of $9.4 to $37.4 billion (or a 0.073% to 0.28% reduction of earnings) for the entire U.S. population due to influenza.

In section 2 I discuss the basics of influenza, how it can affect education, and the existence of cross-type immunity. Section 3 provides the methodology. Section 4 explains the data. Section 5 presents the results, impacts of influenza aggregated across childhood, potential mechanisms, robustness, and policy implications. Finally, section 6 concludes.

2. The Potential Impacts of Influenza on Educational Attainment

Influenza has several unique properties that indicate how exposure to the virus potentially impacts education. First, influenza is notorious for its rapid and unpredictable antibody generator

(antigen) change—the virus is able to disguise itself such that the immune system cannot easily detect it (Cox, Fukuda, 1998, Monto, 2008, Wilschut et al., 2006). Second, today's influenza strain is often the same as an earlier strain (Thompson et al., 2004, Wilschut et al., 2006). These imply that there is considerable variation in influenza and that immunity is possible but not necessarily complete (Smith et al., 2004, Fiore et al., 2010, Glezen, Couch, 1997).

Influenza can have a lasting impact on education. It is a severe childhood disease that is worse for children than typical adults (Beigel, 2008, Cox, Fukuda, 1998, Lemaitre, Carrat, 2010, Wilschut et al., 2006, Evans, Kaslow, 1997).[2] Almond (2006) found large and lasting impacts of *in utero* influenza exposure on completing a high school degree. Children can be symptomatic for up to 10 days due to a single incidence of influenza. Lemaitre & Carrat (2010) implies that children are sick in 1.8 to 7.1 years of childhood (possibly being sick multiple times in the same year). Hospitalization, pneumonia, and other complications can result (including the flare up of existing conditions), which will extend the time ill. Assuming that kids on average missed 10 school days over their primary and secondary education, they would miss 0.34% of their total school days.[3] For moderate levels of influenza exposure throughout childhood, I find that the probability of graduating from high school is reduced by 0.32%. There may be channels other than missing school through which influenza exposure can have a lasting impacting on schooling: Children may not do as well if they are in school when sick. School age children, 7–14 have higher infection rates than adults and are prime influenza vectors (Beigel, 2008, Cox, Fukuda, 1998, Simonsen, 2001). Children's own morbidity can therefore have negative spillovers on their peers, their teachers, and their parents. These negative spillovers can reduce the quality of education received and can reduce consumption, perhaps due to increased parental work absenteeism to care for the sick children (Molinari et al., 2007, Monto, 2008, Wilschut et al., 2006, Oswald, Ridenhour, 2013).[4] The timing of the flu season may also play a role. Influenza generally peaks in January or February, which is about when most schools begin and end terms. Absences or illness-driven poor performance at this time may be especially deleterious. Additionally, remnants of the virus have been found in the heart, brain, and lungs indicating that the virus directly damaged those organs (Beigel, 2008, Simonsen, 2001). This damage may be permanent, reducing future cognitive ability and affecting subsequent academic

[2] Influenza is not to be confused with gastroenteritis, a.k.a. stomach flu, 24-hour flu, or food poisoning (despite these misnomers), or the common cold.

[3] Mikulecky (2013) reports that modal school year length is 180 days. There are 13 years in primary and secondary education. $\frac{10}{180*13} = 0.34\%$ The median child will be sick 4.4 times (Table 5), and the average number of days missed is ~2.5 (Figure 4).

[4] For example, Noble (2012) finds that malaria in children reduces consumption.

performance. Finally, influenza exposure may affect identity formation in an Akerlof & Kranton (2000) sense.[5] Testing and exploring these potential explanations for why influenza exposure can have a lasting impact on educational attainment is beyond the scope of this paper. Nonetheless these provide support for flu exposure having lasting impacts.

Given the nature of the flu and the existence of cross-type (heterotypic and heterosubtypic) immunity in humans (Epstein, 2006, Glezen, Couch, 1997, Boon et al., 2004, Lavenu et al., 2004), influenza providing a long-run benefit is a logical result.[6] A high degree of serial correlation in strains allows for immunity acquisition (Wilschut et al., 2006, Smith et al., 2004, Russel et al., 2008). While the virus can mutate at any point in time, it rarely changes enough to be drastically different from year-to-year (Smith et al., 2004). The dearth of influenza pandemics is evidence of this (Wilschut et al., 2006, Smith, 2006). Additionally, antigenic studies have shown less variation than genetic studies would indicate, and that strains evolve out of a common viral ancestor (Smith, 2006, Russel et al., 2008, Smith et al., 2004). The lack of substantial change in the virus and hetero(sub)typic immunity allows for the acquisition of immunity to play a role in educational attainment. Exposure at earlier ages can then protect children when they are subsequently exposed to the same (or similar enough) strain later in childhood during schooling.

The attributes of influenza suggest that it can have both positive and negative effects on educational attainment, and that effects may vary by age at exposure which can be modeled formally. Consider a simple input/output model of influenza's impact on schooling:

$$(1)\ Education_i = Influenza_i\theta$$

where, Influenza_i is a 1 x A vector for person i measuring exposure to influenza where A represents the maximum age considered during childhood plus one. θ is an A x 1 vector measuring the impact of influenza exposure at a given age on educational attainment.[7]

Immunity acquisition (and cross-type immunity) allows for the seemingly counterintuitive case of $\theta_a > 0$.[8] To better illustrate this, consider the following thought

[5] For example, a child gets sick with the flu and does poorly. They then falsely attribute that poor performance to their own ability rather than the disease. Subsequent notions of identity are then based off of this false assumption, affecting future performance.

[6] Heterotypic, or cross-type, immunity is immunity, for example, to influenza B acquired through exposure to influenza A. Heterosubtypic, or cross-subtype, immunity is immunity, for example, to A-H1N1 acquired by exposure to A-H3N2. Cross-type immunity will be used to refer to both heterotypic and heterosubtypic immunity.

[7] Where $a \in [0, A-1]$ indexes the age at exposure, such that, for example, θ_0 represents the effect of influenza during infancy.

[8] There are two potential mechanisms for this. The first is discussed in text above. The other, and a concern, is culling. Given the data being used, individuals must have lived long enough to be in the sample. Therefore it is possible, albeit unlikely, that a positive effect could result due to the less healthy and less educated dying from influenza. Since the deceased children do not appear in the data (due to their influenza deaths), there is the potential that this could drive a positive relationship. However,

experiment. Assume that immunity is complete and permanent once acquired solely through exposure and assume further that morbidity from exposure during school reduces education. Consider two otherwise identical kids: Child E is exposed during pre-school years (e.g., age 4) while U is unexposed during pre-school. We then expose both children to flu at age 11 when the child is in the 5th grade. E would be unaffected by exposure in the 5th grade while U would be affected, and therefore have worse educational attainment than E. So, for U, $\theta_{11} < 0$ and for E, $\theta_{11} = 0$. If E had not been exposed at age 4, the flu at age 11 would be a detriment for both children. But U has less education than does E, because U became sick with the flu at age 11. This implies for E, $\theta_4 > 0$, and for U, $\theta_4 = 0$, since in the absence of morbidity at age 4, E would have had less education. It was E's early exposure to influenza that resulted in the relative increase in education.

The previous thought experiment is altered by multiple strains, which changes how education can benefit from immunity acquisition. This is discussed formerly in appendix section B. Immunity can be viewed as a continuum. A child exposed early on to one version of the virus has immunity to that version, and to a lesser extent other versions through hetero(sub)typic immunity, more evidence in favor of this is presented in appendix section C (Epstein, 2006, Glezen, Couch, 1997, Lavenu et al., 2004). This could range from not presenting with any symptoms to less severe or fewer symptoms to reduced duration. The presence of multiple flu strains potentially diminishes the benefit of immunity acquisition, but hetero(sub)typic immunity still allows for it (Epstein, 2006, Glezen, Couch, 1997, Lavenu et al., 2004). As the virus becomes less detectable by the human immune system, immunization effects will move to zero and the negative effect of the shock itself will become more pronounced. Additionally, considering multiple strains in the previous thought experiment generates a testable prediction: a child's first exposure to influenza will have a greater impact than subsequent exposures.

3. Methodology

The primary goal of the econometric model is to estimate the effect of influenza exposure throughout the entirety of childhood on educational attainment.[9] Exposure is ideal for use as a measure of the flu's intensity as it does not suffer from simultaneity concerns. It is policy

child death due to influenza is a highly unlikely event, as evidenced by the odds being approximately 1 in 1,750,000. Therefore, culling is not behind beneficial influenza. Were I to add those children that died from influenza back to the sample, they fail to affect the results presented in section 5.

[9] Using exposure (rather than morbidity) lends itself towards an intention-to-treat (ITT) interpretation for β.

relevant, since exposure is the target of the Centers for Disease Control (CDC) through vaccination. Additionally, exposure captures more than own morbidity in that it reflects the overall experience of children due to influenza (e.g. siblings' and classmates' morbidity). I modify (1) to its empirical analogue for estimation:

$$(2)\ Outcome_{isc} = X_{isc}\phi + Influenza_{sc}\beta + \mu_s + \lambda_c + \epsilon_{isc}$$

where $Outcome_{isc}$ represents education and health outcomes for individual i in birth state s and birth cohort c. Specific outcomes are discussed in greater detail in section 4 and presented in Table 1. $Influenza_{sc}$ represents influenza exposure for a given state, s, and birth cohort, c (having the same structure as in (1)). β is an A x 1 vector, the empirical analogue to θ, measuring the impact of influenza exposure on educational attainment. μ_s and λ_c capture state and cohort (i.e. year) fixed effects. Finally, X_{isc} is a vector of controls for race and sex.

Equation (2) is estimated using a linear probability model.[10] The source of exogenous variation in the risk of influenza is at the state/year level. Figure 1 presents this variation as the risk of elderly influenza death that cannot be explained by state and year fixed effects from a state-level fixed effects regression.[11] These residuals are plotted against the year of influenza exposure, with observations labeled by their state. Were state and year fixed effects able to completely explain the influenza proxy then Figure 1 would be a straight line at zero. As can be seen, this is not the case. The residuals presented in Figure 1 is the variation that I exploit to identify the impact of exposure to influenza on educational attainment. Following Bertrand, Duflo, & Mullainathan (2004), I cluster the standard errors at the state to avoid overstating significance that can erroneously result do to serial correlation.[12]

4. Data

The data needs suggested by the econometric model are large. I need a measure of influenza exposure that is available for at least 19 consecutive years for a given birth cohort. The measure must also vary across geography to help separate cohort effects from flu exposure. Data on morbidity, e.g. hospital admissions or doctor visits, is not available with geographic identifiers at an annual frequency for the requisite amount of time. For example, readily available datasets like the Agency for Healthcare Research and Quality's (AHRQ) Healthcare

[10] This is a reduced form examination of Cunha & Heckman's (2007) sensitive and critical periods at an annual frequency.
[11] The regression equation is: $Influenza_{sc} = \mu_s + \lambda_c + \epsilon_{sc}$. The data is the MCOD, see section 4, aggregated to the state-year.
[12] This follows Moulton's (1990) observation that failure to account for cross-correlation in errors due to commonalities inherent in assigning state level variables to micro observations can lead to spurious regressions in an OLS setting.

Cost & Utilization Projects (HCUP) are only reliably available for a more recent time period and a limited sample of states which precludes examining health during the entirety of childhood on educational attainment.

Influenza virulence and exposure is constructed using the National Vitality Statistics' Multiple Cause of Death (MCOD) files, available from 1968 onwards.[13] MCOD is available for well over 19 consecutive years making it ideal for use and also allowing multiple cohorts to be studied. It is a census of all deaths in the United States, providing the underlying cause of death, age at death, state of residence, state of death, and year of death. From this the number of individuals that have died due to influenza in each state-year for a given age group is known.

State-cohort influenza exposure will be used. The identifying variation is from differences in influenza strength (virulence), density, and or fatality in a given state and year. Individual influenza morbidity would be superior in terms of assignment of influenza to individuals, but suffers from selection bias. For example, were hospital admissions-type data to be used, only those individuals with influenza severe enough to warrant hospitalization would be captured. If US data were used, there would be concerns of access to health care as well. By aggregating the data to the state-cohort level, I avoid selection concerns. Everyone in a state-cohort is assigned to the same level of influenza exposure, but attenuation will likely result. From this the mean effect of influenza on the population can be determined. In order to estimate the effect of influenza on later-in-life outcomes, the outcomes data needs to be linkable to individuals at the state-year level.

Mortality will be used as a proxy measure of influenza virulence. Epidemiological and medical studies have found that influenza mortality is predictive of morbidity (Monto, 2008, Simonsen, 2001, Wilschut et al., 2006). Mortality is one of four methods used to track epidemics in the US and Europe (Monto, 2008, Simonsen, 2001, Wilschut et al., 2006).[14] Mortality was superior in predicting the actual number of influenza cases during several epidemics (Cox et al., 2001, Domínguez et al., 1996).

Those who are most at risk to influenza are children, the chronically ill, and elderly. Of those at risk to influenza, it is the elderly who represent the lion's share of deaths (Cox, Fukuda, 1998, Monto, 2008, Wilschut et al., 2006). This can be seen in Figure 2, which presents all influenza deaths by age at death from 1968–2000. The vertical lines represent the end of

[13] This data can be accessed online at: http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm

[14] The other three being: (1) collaborating laboratories, which track specimens positive for the influenza virus; (2) state and territorial epidemiologists, which estimate the number of cases of influenza in their jurisdiction; and (3) sentinel physicians, which track and report the number of visits for and patients with influenza.

childhood, at age 18, and the beginning of seniorhood, at age 65. Childhood influenza death is a very uncommon event (Simonsen, 2001). Between 1968 and 2000 there were ~1 million child deaths, of which 1,256 were due to influenza. During this time the average child population was ~70 million, therefore, the probability a child dies from influenza is ≈ 0.000006 in any given year. To put this in perspective, a child is 400 times more likely to be struck by lightning during their childhood than die from influenza.[15]

While it is almost exclusively the old who die from influenza, it is primarily the young who get sick from influenza (Beigel, 2008, Cox, Fukuda, 1998, Wilschut et al., 2006). Influenza attack rates are highest for children, ranging from 10% in non-epidemic years to 40% in epidemic years (Cox, Fukuda, 1998, Monto, 2008). Due to the rarity of child deaths and the positive correlation between elderly deaths and child morbidity, this paper will use elderly death rates as the measure of influenza strength in a given state-year.[16]

Elderly influenza deaths are correlated with child influenza morbidity. I know of no data which presents morbidity data by age for 1969–2000 inclusive. I have, however, found data that covers the time period with proxies for child morbidity. Figure 3 is a monthly national time-series from the MCOD showing both elderly influenza deaths (left y-axis) and child influenza deaths (right y-axis). As can be seen, the two peak and trough at near identical times (evidenced by a correlation of 0.7915), indicating that there is one flu season, and it is the same for the elderly and children. The flu season is October to April; it peaks around January, which can also be seen in Figure 3 (Beigel, 2008, Cox, Fukuda, 1998, Wilschut et al., 2006). Figure 4 presents the quarterly elderly influenza death rate and the number of days that children spent sick in bed (bed days) as a percent of the days possible.[17] Bed days come from the National Health Interview Survey (NHIS) from 1969 to 2000 (Minnesota Population Center, State Health Access Data Assistance Center, 2012).[18] The NHIS is aggregated to the national level for each quarter. Bed days are the number of days spent in bed during the past two weeks until 1997, then bed days is for the previous year. These two definitions were combined into a single variable and made to be percentages. The correlation between bed days and elderly flu deaths at a national quarterly level is 0.6343. Bed days are the number of days sick in bed for any condition, not just

[15] This information is provided by the National Weather Service online at: http://www.lightningsafety.noaa.gov/medical.htm
[16] Additionally, using child deaths would be more susceptible to culling, see section 3. Deaths have been previously used as a proxy for virulence, e.g. Almond (2006) or Evans et al. (2012).
[17] By definition influenza lasts for at least 72 hours, and can last as long as a week or 10 days. Afflicted children will not only physically miss school by being sick in bed, they can also retain less of what is taught during school while physically present due to morbidity. Exposure as defined is not just limited to time spent sick in bed.
[18] The NHIS is accessible online at: http://www.ihis.us.

influenza. However, children are sick in bed when the elderly are dying from influenza. Both Figure 3 and Figure 4 provide evidence that elderly influenza deaths are correlated (and therefore represent a good proxy) with child morbidity.

Using total flu deaths at the state level would fail to account for state populations. Data on state populations is taken from the Survey of Epidemiology and End Reports (SEER).[19] This is available from 1969 onwards. Specifically, the elderly population for each state-year is used. Equation (3) explicitly provides influenza exposure/virulence to be used:

$$(3)\ Influenza_{isc} = \frac{elderly\ influenza\ deaths_{sc}}{elderly\ population_{sc}} \in [0,1)$$

where i indexes individuals. Influenza exposure is constructed at the state-cohort level, where s indexes an individual's birth state and c indexes the birth cohort. Effectively, this is the probability that an elderly person dies of influenza in a given state-year.

Influenza is plausibly exogenous. The strength, virulence, fatality, etc. of the virus is determined by random mutations in the virus's antigenic structure; not much more is known about how the virus evolves and is still an active area of research (Denoued et al., 2007, Russel et al., 2008, Smith et al., 2004, Smith, 2006).[20] Influenza exposure as defined in (3) does not suffer from simultaneity concerns when considering educational attainment. By considering influenza as assigned in childhood (by birth state) on an adult's educational attainment, mechanically there cannot be a simultaneity issue, flu exposure happened first. Exposure at the state-year level is like intention-to-treat (where morbidity would be treatment-on-the-treated). In which case, exposure is not going to be influenced by the child's innate health status, and therefore avoids the simultaneity concerns that morbidity would suffer from. By using elderly influenza deaths, rather than using children's actual influenza experience (either morbidity or mortality), there exists no mechanical relationship where one of the individuals in the study is also represented in the influenza proxy.

Outcome and control variables are obtained from the 1980–2000 US Censuses, IPUMS version 5.0 (Ruggles et al., 2010). The US Census is ideal for analysis in this context. It provides a large sample size, which can detect an anticipated small and noisy effect of influenza exposure in any given year of childhood.[21] It allows for the examination of outcomes *ex post*

[19] The SEER data is available online at: http://seer.cancer.gov/popdata/download.html.
[20] The most that is known is that influenza is more easily vectored in colder, more arid conditions, but why and how this is the case is still not well understood(Denoued et al., 2007, Lamb, Brannin, 1919). General climate characteristics will be captured by state and year fixed effects, thereby controlling for this potential concern.
[21] While influenza is a severe childhood disease, the CDC did not recommend that school age children be vaccinated until 2006. Additionally, extreme negative outcomes like death are rare in childhood. State-cohort assignment is going to be imprecise, with

(i.e. after schooling has been completed), avoiding simultaneity concerns. And the Census collects data on state of residence, state of birth, and birth year, allowing it to be combined with influenza exposure at the birth state–birth cohort level.[22]

The primary outcomes considered are whether or not at least a (i) high school diploma (or equivalent), (ii) some college, or (iii) a bachelor's degree was obtained. For these education variables, the analysis uses the 2000 Census for individuals born between 1969 and 1977. Summary statistics for these individuals are presented in Table 1. Table 1 is discussed in greater detail at the end of this section.

In addition to the previous education variables that look at completed schooling, this paper also explores the potential for education as a mechanism by exploring how influenza affects on-going education. This is done using Oreopoulos et al. (2006) "grade-for-age" measure (GfA) and attending post-secondary education. GfA is equal to one if an individual is at or beyond the modal level of schooling for all individuals with the same birth year, (birth quarter), and birth state. GfA is considered for both lower/mid primary school (ages 7–11) and high school (ages 15–19).[23] For the primary school sample only the 1980 Census is used for individuals born between 1969–1973, to best match Oreopoulos et al. (2006). For the high school sample the 1990 and 2000 Censuses are used for individuals born between 1970–1974 and 1980–1984. Birth quarter cannot be used for this sample as these Censuses lack it. The attending post-secondary variables are attending college and attending graduate school.

In addition to education variables and control variables, Table 1 also presents summary statistics on the few health related variables available in the 2000 Census. Those presented are whether or not the individual has any disability that presents a difficulty for working (the 2000 disability definition), and whether or not there are cognitive, ambulatory, and seeing or hearing difficulties. These variables are presented as equal to one if there is a disability or difficulty. These health-related variables may point to a potential mechanism through which influenza can

---

attenuation a likely result. Hence, if there is an effect of influenza on educational attainment, the expectation is that that effect will be small. These motivate a data source with a very large sample size.

[22] Individuals in the Census are matched to their appropriate influenza exposure using their birth year, $Birth\ year = Census - (Age + 1)$, and state of birth. For individuals who reside in their birth state at the time of the Census, they are assigned a complete history of influenza exposure. Therefore, individuals who have moved are omitted from the analyses when looking beyond the impact of health in the birth year.

[23] The only way to examine GfA is to consider multiple birth cohorts simultaneously (i.e. multiple grades), due to multicollinearity between the influenza variables and the state fixed effect. Consider (2) modified to consider a single birth cohort, 1970, for one of the GfA outcomes: $GfA_{is1970} = X_{is1970}\phi + Influenza_{s1970}\beta + \mu_s + \lambda_{1970} + \epsilon_{is1970}$. λ_{1970} will drop out since it is now perfectly collinear with the constant term. $Influenza_{s1970}$ now only varies at the state level, and all of these variables are multicollinear with the other variables represented in $Influenza_{s1970}$ along with the state fixed effect, μ_s. Hence GfA cannot be examined for a single cohort.

affect education. Remnants of influenza have been found in major organs, e.g. the brain, indicating that the virus damaged those organs and reduced healthy function (Beigel, 2008, Simonsen, 2001). Hence, working, cognitive, ambulatory, and seeing/hearing difficulties may result from influenza exposure (as the influenza virus could directly and irreparably damage cells/organs responsible for those functions) during childhood and be a pathway through which education is reduced. Of the health variables considered, cognitive difficulties is the best candidate for a mechanism, as this specifically refers to whether the individual has difficulty learning, remembering, or concentrating and since viral remains have been found in the brain.

Table 1 presents summary statistics on all variables discussed above. Summary statistics are presented separately for all individuals, individuals who reside in their birth state (the main sample), and those who have left. There are differences between movers and non-movers. Of those born between 1969 and 1977, 67.5% remain in their birth state. During childhood, non-movers are marginally better off in terms of schooling, but in adulthood the educational advantage goes to movers. There is not a clear health advantage for either group, with non-movers being more likely to report having a cognitive, ambulatory, and or seeing/hearing difficulty, and movers are more likely to report having a disability that makes work difficult. On the whole, movers are more educated, with no other major differences existing between the two groups. Given that this study is largely conducted on a less educated population, there is the potential for influenza exposure's impact on education to be understated.

5. Results

Complete childhood exposure results for the main outcomes, education, are presented as Figure 5 to Figure 7, and in Table 2. The figures present the percentage change in education due to exposure to epidemic influenza, γ_a, for each age of exposure during childhood, with 95% confidence interval shown as dot-dash.[24] The presentation of the results stem from (2), but are not the coefficients themselves:

$$\gamma_a \equiv \frac{(Q_3 - Q_1)_a \times \beta_a}{\overline{outcome}}$$

[24] Note that Table 2 and Figures 5 to 9 (along with all other tables except Tables 9 to 11) do not present the coefficients from (2).

where Q_i represents the i[th] quartile.[25] β_a is the coefficient corresponding with the a[th] age of exposure. And $\overline{outcome}$ is the sample mean, listed at the top of each column in Table 2. Multiplying coefficients by the interquartile range simulates the *experiment* of exposing one group of people to an innocuous form of influenza as a *control* group, e.g. seasonal flu, and another to a more virulent form as the *treatment* group, e.g. epidemic influenza. Therefore each parameter in Table 2, or corresponding figure, is the effect of being exposed to highly virulent influenza. Each of these is then divided by the mean of the dependent variable to provide a percentage interpretation. For example, being exposed to epidemic level influenza in the sixth grade reduces the probability of obtaining at least a bachelor's degree by 5.15%. Since the coefficients are not presented, p-values are presented instead of standard errors. As mentioned in section 3, these p-values are taken from standard errors that allow for within-state correlation in errors.

Before school (infancy to age five) there is generally a positive and statistically significant finding of 0.05%–3.72%. Epidemic influenza (statistically insignificantly) increases the probability of obtaining a high school diploma by 0.05–0.14 percent in any given year prior to school.[26] There are greater effects found for completing at least some college, with greater virulence (statistically significantly) increasing the probability of at least some college by 0.22%–0.57% in a given year. The largest coefficients and most statistically significant effects are found for a bachelor's degree. Harsher influenza prior to schooling increases the probability of obtaining a bachelor's degree by 2.56%–3.72%. In the early years of schooling, ages six to eight, the positive effect persists, but statistical significance drops.

These positive parameters support an immunity acquisition story. However, the immunity acquisition mechanism itself cannot be tested (due to data limitations), the next sub-section, 5.1, will examine other potential mechanisms.

It is in upper primary school, beginning with third grade, that the result becomes negative. This negative finding persists for the remainder of primary school and softens entering secondary schooling. During this time influenza exposure serves to reduce educational attainment by 0.21%–5.15%. Note that the magnitudes of these are generally greater than those experienced prior to schooling. Influenza has the greatest impact on obtaining a bachelor's

[25] Note that $Q_3 - Q_1 \geq 0$, as such the effect will have the same sign as the coefficient, with $\beta_a > (Q_3 - Q_1)_a \times \beta_a$. Also note that the interquartile range in this case is approximately equal to the standard deviation, this interpretation can also be applied.
[26] Higher levels of influenza will always refer to the difference between influenza virulence at the 75[th] percentile and influenza virulence at the 25[th] percentile. This is meant to be synonymous with epidemic influenza, greater/more virulent influenza, harsher influenza, or simply in reference to exposure.

degree (Figure 7). Virulent influenza reduces the probability of a high school diploma (Figure 5) by 0.21%–0.53% in a given year during upper primary school. For at least some college (Figure 6), moving from the first to the third quartile of exposure reduces reports by 0.73–1.39 percent. And epidemic influenza decreases the probability of obtaining at least a bachelor's degree by 0.86%–5.15% between the ages of nine and twelve.

In secondary school there continues to be a negative effect of increased viral strength, but these are dwarfed in magnitude by those for earlier grades and most are statistically insignificant. There are several noteworthy exceptions. The first being junior high school exposure to higher levels of influenza reducing the probability of obtaining a bachelor's degree by 2.70%–5.49% (statistically insignificant for the other two educational outcomes). Second, is the statistically significant (for high school only) positive bump for flu exposure during the penultimate year of high school (age 17). Finally, there is the senior year in high school (age 18). Virulent influenza at this time decreases the probability of education by 0.25%–1.33% and is statistically significant for all outcomes.

There is a clear dichotomy for influenza's impact on educational attainment. Before entering school children benefit from exposure and during school exposure is harmful.[27] To analyze how timing of health insults affects education we can consider the magnitude and significance of a given year's effect. Generalizing across the education outcomes, 3rd–6th and 12th grade play important roles for educational attainment and are times when exposure to influenza should be reduced or otherwise mitigated, but there is an important beneficial role of exposure prior to entering school.[28] This provides a partial answer to how and when timing matters, but it fails to take into account when children are likely to experience a shock. Sub-section 5.2 tackles this problem to provide a complete answer by considering when shocks are most likely to occur and conditional on a shock occurring the impact of that shock at that time.

5.1. Mechanisms

The question I turn to now is what pathway is driving these results. Two mechanisms that the Census data allow exploration into are especially relevant: concurrent education and

[27] This has also been tested using a pre and post entering school design, $Outcome_{isc} = X_{isc}\phi + pre_{sc}\beta_{pre} + post_{sc}\beta_{post} + \mu_s + \lambda_c + \epsilon_{isc}$. Where the pre period is ages zero to five, and the post period is all subsequent ages. Across no outcome is β_{post} statistically significant, although it is negative. β_{pre}, however, is statistically significant and positive. In no case can the pattern (of negative before entering school and negative after entering) observed through estimation of (2) be rejected.

[28] In the terminology of Cunha & Heckman's (2007), upper-primary school is a critical period and the senior year in high school is a sensitive period for influenza impacting educational attainment; before school is a sensitive period for obtaining immunity.

health. For concurrent education, if influenza plays an important role for schooling then it may be detectable during school. For example, if exposure to influenza increases the number of days missed in primary school it may decrease the probability of on-time progression through school (as measured by GfA). For health, the virus is reported to have damaged vital organs, including the brain, this can then manifest as an increase in reports of cognitive difficulties (Beigel, 2008, Simonsen, 2001).

I consider two different types of incomplete education: Oreopoulos et al.'s (2006) GfA and pursuing higher education (Figure 8).[29] GfA is an imprecise measure of on-time progression; that imprecision is shown by large standard errors.[30] Unfortunately due to this, concurrent education cannot be conclusively ruled out as a potential mechanism, but the GfA results provide minimal support for it. At best, there is mild evidence that influenza reduces on-time progression during early primary school.

I further test on-going education as a mechanism by examining how attending post-secondary (Figure 8) education is affected. These results generally confirm the educational attainment findings, as similar patterns are observed. There are increases prior to entering school, 2.08%–7.61%, and decreases mostly during upper primary school, 4.95–22.32 percent. The college attendance regression dominates all others in the magnitudes of its effects, generally one order of magnitude larger, and may be cause for alarm. The results from the graduate school attendance regression are less statistically significant than others and have magnitudes that are one order of magnitude less than the others. This can be seen as an out-of-sample replication of the main result and as pointing to on-going education as a potential mechanism. In total, there is mild evidence that concurrent education, e.g. missing school, is a pathway through which completed education is reduced.

The Census data allows for examination into several health-related characteristics: work disability status and whether or not there was cognitive (Figure 9), ambulatory, or seeing/hearing difficulties. These results are presented in Table 4, which follows the same structure as Table 2. All of these variables are defined as equal to one if there is no disability or difficulty.[31] These are measures of permanent health. The health measure in the Census that is the most likely pathway for influenza to affect education is cognitive difficulties (see section 4). Moreover, the greatest magnitude of these effects is for cognitive difficulties, with greater viral potency

[29] Tabular form of the results is presented in Table 3. The figure for attending graduate school and GfAs have been excluded.
[30] Failing to be at one's GfA has a potentially ambiguous effect on educational attainment itself, refer to Oreopoulos et al. (2006).
[31] Such that, like the education outcomes, positive results represent good outcomes and negative ones bad.

increasing the probability of cognitive difficulties by 0.09%–0.22% when exposed prior to schooling. Influenza has no effect on the remaining notions of health. There are virtually no positive (i.e. beneficial) and statistically significant effects of influenza on reports of health difficulties. That is, when influenza does have an effect it reduces the permanent stock of health. For ages 7 to 14, the time epidemiologists have identified as the most likely for influenza morbidity, most of the impacts are statistically insignificant. The key exception is cognitive difficulty. During junior high school is when completing college is most hurt by influenza. Also during junior high epidemic influenza increases cognitive difficulty by 0.11%–0.17%. This is evidence that cognitive function may be a causal pathway through which influenza can affect education.

5.2. Aggregate Effects

Before the question of when and how timing matters was answered partially by looking at magnitude and significance. This is only a partial answer because each of the results presented is the effect of a shock given that a shock occurred that year. The total effect must account for when and how many shocks are likely to occur.[32] For example, while there may be a dramatic effect of shocks in the 3rd–6th grade range, if children are unlikely to experience influenza at this time the importance of these shocks is diminished. However, this is, in fact, the most likely time for children to experience influenza (Beigel, 2008, Cox, Fukuda, 1998, Simonsen, 2001).

I obtain the probability of a shock in a given year using the probability of influenza morbidity. Lemaitre & Carrat (2010) report the relative illness ratio (RIR) of influenza at each age, where:

$$RIR_a = \frac{sick_a / \sum_a sick_a}{N_a / \sum_a N_a}$$

That is, the RIR is the proportion sick at a given age divided by the proportion of the population at that age. This provides a relative notion of morbidity, but is not the probability of experiencing a shock at a given age. To obtain this from Lemaitre & Carrat's (2010) RIR the

[32] Other notions that could be considered for a total effect are mean effect, $\sum_a \frac{\gamma_a}{A}$, and sum of effects, $\sum_a \gamma_a$. However, mean effect would force an equal weight across all years and force the number of times ill with flu to be unity (most generously the odds of this being 1 in 50). Summing effects would mathematically require that a shock was experienced in each and every year during childhood, i.e. the effect of A, e.g. 19, influenza spells, this is a highly unlikely event, the odds (generously) being 1 in 181 million. Hence neither mean nor sum of effects would be appropriate. The notion of total considered is also the most likely, with (most conservatively) 84% of children getting sick with the flu at least once.

proportion of the population that experiences morbidity is needed. The proportion of the population that experiences influenza is a parameter, $\eta \equiv \frac{\sum_a sick_a}{\sum_a N_a}$.

$$RIR_a \times \eta = \frac{sick_a}{N_a} = Pr[sick\ with\ influenza \mid age = a] \equiv \rho_a$$

ρ_a is simply the probability that a child experiences a shock at a given age. Cox & Fukuda (1998) and Monto (2008) report $\eta \in [0.05, 0.20]$, which in turn bounds ρ_a. Table 5 shows the relationship between η, the child morbidity rate ($\sum_{a=0}^{A-1} \frac{\rho_a}{A}$), and the mean number of times that a child will be sick during their childhood ($\sum_{a=0}^{A-1} \rho_a$) for three levels of influenza virulence: mild, median, and epidemic. For example, in a given year, if influenza has a median attack rate, then 12.5% of the population gets sick with the flu, implying 23.2% of children would get sick, and were this level constant during childhood the mean number of times that children would be ill in their childhood is 4.4. The range of morbidity during childhood is 9.3%–37.2%, which is the range reported by Cox & Fukuda (1998) and Monto (2008) for children. The mechanical relationship provides a nice interpretation: the aggregated effect can be tied to the number of severe shocks experienced during childhood, or the number of times that a child is sick with flu. This allows for the construction of an aggregate, or total, effect of influenza during childhood. Both the impact of the shocks and the probability that those shocks occur are accounted for in this aggregate effect.

The overall impact of influenza on education across childhood is presented in Table 6. The rows for this table are the same as those for Table 5: mild, median, and epidemic flu seasons throughout childhood. In Table 6, I sum across all years in childhood the probability of experiencing a shock multiplied by the effect of that shock, $\sum_{a=0}^{A-1} \rho_a \gamma_a$.[33] Table 6 also presents standard errors for the aggregated effects.[34] Note that all aggregated effects are highly statistically significant. For the main educational attainment outcomes there is a negative effect. Education is reduced by influenza. For educational attainment, where the positive prior-to-school and negative after-school was most pronounced, the negative effect dominates, implying that 3rd–6th and 12th grade are indeed the most important. Individuals are 0.13%–0.52%,[35]

[33] Were an ITT interpretation applied to β_a or γ_a, then the aggregate effect could be thought of as a treatment-on-the-treated (TOT), where ρ_a can also be thought of as the proportion of the relevant population with influenza (i.e. treatment).

[34] Standard errors were calculated as: $\sqrt{\sum_{a=0}^{A-1} \rho_a^2\, var(\gamma_a) + 2 \sum_{a<b} \rho_a \rho_b cov(\gamma_a, \gamma_b)}$, where $var(\gamma_a) = \left(\frac{(Q_3 - Q_1)_a}{outcome}\right)^2 \times var(\beta_a)$, and $cov(\gamma_a, \gamma_b) = \left(\frac{(Q_3 - Q_1)_a}{outcome}\right)\left(\frac{(Q_3 - Q_1)_b}{outcome}\right) \times cov(\beta_a, \beta_b)$.

[35] Almond (2006) finds that males exposed to the 1918 Spanish Flu pandemic were 1.4 to 2.1 percent less likely to obtain a high school diploma. This finding is approximately two-fifths the size of Almond's (2006) estimate.

0.39%–1.56%, and 0.74%–2.96% less likely to report obtaining at least a high school diploma, some college, or a bachelor's degree, respectively, due to influenza exposure during their childhoods conditional on mean morbidity. It is important to point out that if children were exposed to median levels of influenza (Table 5) and therefore missed the average number of days (Figure 4), they would miss 0.34% of their primary and secondary education, the corresponding impact I find is 0.32% (for moderate influenza exposure on the probability of graduating from high school).

Table 8 presents the total effect for adult health.[36] An implication is that influenza during childhood reduces the permanent stock of health. The greatest of these is for cognitive difficulties (the most likely candidate as a health mechanism), with influenza exposure increasing reports by 0.08%–0.35%. For the remaining adult health outcomes, work, ambulatory, or seeing/hearing difficulties, the total effect of influenza during childhood serves to increase the probability of incidence by 0.02%–1.62%. When the total effect of influenza is considered, children are made worse off by increased exposure and morbidity. Despite potential immunity acquisition gains, influenza is detrimental to education and health.

5.3. Impacts of Influenza Around Birth

In the literature there has been considerable attention paid to health *in utero* or during infancy (for example, see Currie (2011)). Given this attention and that influenza is a mild shock to health, examination in the effect of influenza during infancy is warranted. The un-modified coefficients, i.e. β_0, are presented in Table 9 for the education outcomes and in Table 10 and Table 11 for education and health mechanisms, respectively. Influenza does not have an effect during infancy. These coefficients are statistically insignificant and are close to zero, with magnitudes ranging from 0.04–0.17 percentage point changes at the mean. These are smaller than those found for the birth year in conjunction with the remainder of childhood. The finding of a null effect may be because mothers can pass on their own antibodies, and therefore immunity to influenza (Wilschut et al., 2006). In regards to health or education, there is also little to no effect of influenza during infancy.

[36] Table 7 is presented for completeness. However, since little was learned from the regressions (due to imprecision) used to create it, it is not discussed here.

5.4. Importance of the First Exposure

A testable hypothesis of the thought experiment with multiple strains was that children's first exposure to influenza should have the greatest impact (discussed in length in appendix section B). To test this the econometric framework needs to be modified, specifically the notion of influenza. In this case a "shock" is needed. A shock is defined as one that exceeds the third quartile, where:

$$AS_{sc} = 1\big(Influenza_{sc} \geq Q_3(Influenza_{sc})\big)$$

To examine the effect of the first shock (FS) consider the ath element (note that for a = 0, FS is the same as AS):

$$FS_{sca} = 1(Influenza_{sca} \geq Q_3(Influenza_{sca}) \cap Influenza_{scb} < Q_3(Influenza_{scb}) \; \forall \; b < a)$$

where a indexes the age of influenza exposure and b therefore represents some previous year's exposure. Then modify (2) with $Influenza_{sc} = \{AS_{sc}, FS_{sc}\}$. The estimation can then proceed as before with the linear probability model. For ease of exposition, Table 12 presents the total effect of first versus any shock for education, Table 13 does so for education mechanisms, and Table 14 does so for health (analogous to Table 6, but only presenting median influenza). The latest any individual experiences a first shock (conditional on ever experiencing one) is at age eight. Hence, for each outcome two aggregated effects are presented: one for the entire childhood and one for only the first eight years. What is clear from this is that first shock does indeed have a greater impact than subsequent shocks, as evidenced by its sum being greater than that for any shock. This is true for both total effects considered, eliminating the possibility that a greater available time span for any shock could be reducing its weighted sum. This also holds on an individual coefficient basis, where 85% of the first shocks are greater than their corresponding any shock. And the standard errors indicate that for all outcomes confidence intervals are not overlapping, hence the testable prediction that the first shock is more important is verified.[37]

5.5. Robustness

The next thing to consider is the robustness of these results. The educational attainment results were to an extent reproduced in two of the on-going education mechanisms, confirming the pattern of sign and statistical significance, and confirming the sign of the overall effect. Additionally, this paper considered a different clustering structure, state-birth-cohort. This de-

[37] Alternatively, shocks were defined as being greater than or equal to the 50th and 90th percentiles. This failed to impact the conclusion that the first shock is the most important.

emphasizes the serial correlation within states, but allows for correlations within a birth cohort. Statistical significance increases in most instances. Probit models were also employed, with the magnitude slightly greater and the same pattern of sign and significance. The results are also robust to culling.[38]

Of further concern is spurious significance due to greater variation in influenza virulence. The Russian Flu pandemic of 1977 is especially responsible for greater variance in influenza. In general it is the earlier years of influenza 1969–1977 that are responsible for greater variance. The variance is greatly reduced for exposure at ages 10 and beyond. To further investigate the possibility of spurious significance due to increased influenza variance in the early years this paper considers the (absolute value of the) t-statistics plotted against age at exposure, presented in Figure 10. Two lines are shown: The solid-black straight line is from a univariate linear regression where the coefficient is -0.0298 with a t-statistic of -8.80.[39] The gray-dash-dot line tracks the mean t-statistics across all outcomes. Both lines confirm that the earlier years are more likely to find significant effects, which runs counter to the greater variation in earlier years. I have ascertained whether or not this is due to increased variance for those early years or whether earlier ages just matter more. Most outcomes exhibit small changes in t-statistics across changes in variance. The outcomes that are of potential concern for spurious significance are cognitive difficulties, bachelor's degree, attending graduate school, and GfA during high school. For the remaining outcomes there is no concern of spurious significance, as the relationship between variance and significance is non-existent.

The negative finding of influenza exposure during upper primary school is robust to concerns of spurious significance induced by greater variance in influenza. The results for cognitive difficulties and obtaining a bachelor's degree may warrant further scrutiny for the effect of influenza prior to schooling. However, the positive finding prior to schooling for obtaining a high school diploma and attaining at least some college are robust to this consideration.

---

[38] There is the potential for a positive coefficient to result from culling. Specifically, children could have been included in the sample but were not due to influenza deaths and might have had minimal education had they lived. And the worry is that by their exclusion the coefficients are positively biased. The MCOD is a census of all deaths in the US. I therefore know the relevant characteristics of those children that were potentially culled by influenza. This allows me to test for culling by *resurrecting* the eligible children, those born between 1969 and 1977. There are 233 such children. MCOD does not contain the relevant educational outcomes. However, under the concern of culling, these children would have had low levels of education, so I fill in their education with zeros. These culled children are then added back to the sample. There are two ways to add these children to the data: (a) simply add them, N = 723,958, or (b) to maximize the potential for culling, replace children that have a bachelor's with the culled children, N = 723,636. The results are unaffected, in terms of sign, significance, and magnitude, by the inclusion of these children under either scenario. While culling had a very small chance of presenting a problem, this demonstrates that it is of no concern. Culling can therefore be rejected as a potential source of bias.

[39] $tstatistic = flu\ exposure\ age\ \psi + \epsilon$

Despite the preceding discussion of robustness, there may be lingering concerns of omitted variable(s) bias (OVB). For example, there may be something else which is not captured in the state and birth cohort fixed effects that varies in a similar way as influenza exposure rates that could be driving the results. If this were the case, it might be expected that incorrectly assigned influenza exposure would exhibit a similar pattern. I present a falsification strategy that rejects the null of OVB.

Educational attainment being (generally) completed by age 25 is used to incorrectly assign influenza exposure to individuals and falsify the previously presented results. Specifically, I use three different (partially overlapping) cohort sets: (1) those 25 years older than individuals included in my sample, (2) 30 years older, and (3) 35 years older. Since education is completed, *future* exposure to influenza for these individuals should have no effect on their educational attainment. However, there are possible exceptions where there could be economically meaningful impacts.[40] Only the educational attainment outcomes are considered for falsification: there are meaningful impacts of influenza for the health outcomes and the education mechanisms already suffer from power problems.

The falsifications provide evidence that the real regressions are detecting something unique to influenza during childhood, rather than OVB.[41] Ideally, the false regressions would result in a flat line at (or statistically indistinguishable from) zero that is statistically different from the real coefficients. This is not entirely the case. Due to large falsification standard errors (as expected) in most cases I cannot reject that the real coefficients are different from the false ones. This is less true for obtaining a bachelor's. The false regressions do not exhibit the same pattern, as do the real regressions; either within a cohort set (across outcomes) or within outcomes (across cohort sets). In the falsifications for high school, there are three to four significant variables. In the real regression, there are seven statistically significant coefficients. This is 1,222 to 7,381 times less likely than the falsifications.[42] In comparison, things look even better for the some college and bachelor's outcomes with the pattern of real significance being 202,864 to 119.2 trillion times less likely than that observed in the falsifications. In

[40] For the fake early years (fake infancy to fake age 3 or so), it might be expected that there are economically (and statistically) real impacts of influenza at this time as individuals may complete educational attainment at later ages. Additionally, there may be economically (and statistically) real impacts in the latter years (fake age 15 onwards), as influenza culling becomes more likely as life expectancy is reached.

[41] The falsifications are available from the author upon request.

[42] The relative likelihood is obtained using a binomial distribution. A success is a statistically significant coefficient with probability 0.05. There are 19 trials, corresponding with the number of coefficients.

combination, this suggests that the real regressions are not suffering from OVB and are portraying something unique to influenza exposure during childhood.

5.6. Lost Wages

The results can be used both to determine the number of lost degrees and the dollar cost of those lost degrees. This is straightforward to do within the sample. Extending it outside of the sample requires that the result found within the random sample holds for the population considered. The Census is a random sample, therefore these results can be extended to the entire US population that was 23 to 31 in 2000. If this is further extendable to the entire US population, then the results imply that influenza reduced the number of high school and college graduates by 311,757–1,247,029 and 431,801–1,727,203, respectively.

Given these lost degrees, the dollar cost can be calculated if the dollar benefit of a given degree is known. The Bureau of Labor Statistics reports the dollar benefit of a given degree.[43] Assume that a lost degree results in the next lowest level of education instead, e.g. not getting a high school diploma implies instead getting less than a school diploma. Lost high school diplomas and bachelor's degrees result in lost wages of $9.4 to $37.4 billion for the US population due to influenza. Alternatively, each birth cohort lost $145 to $582 million due to influenza, with sicker cohorts losing more.

6. Conclusion

Influenza is a severe childhood disease that presents an annual threat to children. Previous studies have found that poor health reduces education, e.g. Currie (2011). My paper extends that finding to a type of health that is unique in that immunity can be acquired by experiencing the negative health shock, allowing for the shock to be beneficial if experienced at the correct time. This paper also extends those findings to a health shock that is constantly a risk and is milder than previous shocks considered, see Almond & Currie (2011). Few studies have comprehensively considered health during childhood, but none have looked when (down to a single year) health matters during childhood. This paper examines health at each year of childhood, answering the question of when in childhood health matters most for education.

[43] Accessible online at http://www.bls.gov/emp/ep_chart_001.htm

Influenza exposure causes reduced education and worse health outcomes. The overall effect of influenza eclipses the dichotomy in timing of exposure that this paper uncovers. Exposure prior to formal education during childhood is found to increase the probability of educational attainment by 0.05–3.72 percent. This gain is possibly due to acquiring immunity (allowed by cross-type immunity) when missing school is not an issue. This is in contrast to exposure after commencing formal education, where an overall detriment is found. During grades three to six (ages 9 to 12) there is a significant negative impact on the outcomes with educational attainment reduced by 0.21%–5.15%. There is also a significant negative effect during the senior year in high school, but smaller in magnitude, 0.25–1.33 percent. These findings are robust to an alternate clustering structure, considerations of spurious significance, falsification, and culling concerns. When the total effect of influenza is considered, children are hurt by flu. Depending on the level of influenza during childhood: a child obtaining (at least) a high school diploma is reduced by 0.13%–0.53%, and the probability that a child obtains (at least) a bachelor's degree is reduced by 0.74 to 2.96 percent due to influenza exposure.

The answer to when timing matters most for health on education is third through sixth and twelfth grade. Poor health reduces education at these times. This is when reductions in schooling are the greatest and these are the ages where children are most likely to get sick with the flu. Epidemiological and medical studies have determined the health benefits to society from vaccinating this population through reduced child morbidity and reduced overall morbidity, as this age group is most likely to spread influenza, see, for example, Sočan et al. (2010). This paper further motivates that call for vaccination by exhibiting the educational costs to children especially at ages nine to twelve who are exposed to severe influenza. The results imply that there were 743,558 to 2,794,232 lost high school diplomas and bachelor's degrees due to influenza exposure. These lost degrees resulted in lost wages of $9.4 to $37.4 billion, with a given birth-cohort losing 145.5 to 582.3 million dollars in wages conditional on the severity of influenza during their childhoods.

The positive findings prior to entering schooling prescribe policy intervention. Children who had greater exposure when school was not at stake were made better off. Hence, if exposure, and therefore immunity, can be increased without school being missed, then children will be made better off with greater educational attainment. Inducing influenza morbidity prior to schooling would be ill advised, despite potential educational advantages, as cognitive difficulties were found to increase resulting from exposure at this time. This is a clear indication

of the gains of vaccinating school-age children against influenza, while also vaccinating younger children. The CDC's current policy of vaccinating everyone greater than six months old is supported by this paper and should have educational, in addition to health, benefits.

A. References

Akerlof, G.A., Kranton, R.E. 2000. Economics and identity The Quarterly Journal of Economics 115, 715-753.

Almond, D. 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, 672-712.

Almond, D., Currie, J. 2011. Killing me softly: The fetal origins hypothesis Journal of Economic Perspectives 25, 153-172.

Baird, S., Hicks, J.H., Kremer, M., Miguel, E. 2011. Worms at work: Long-run impacts of child health gains.

Barreca, A.I. 2010. The long-term economic impact of in utero and postnatal exposure to malaria Journal of Human Resources 45, 865-892.

Beigel, J.H. 2008. Influenza Critical Care Medicine 36, 2660-2666.

Black, S.E., Devereux, P.J., Salvanes, K.G. 2007. From the cradle to the labor market? the effect of birth weight on adult outcomes The Quarterly Journal of Economics 122, 409-439.

Bleakley, H. 2009. Economic effects of childhood exposure to tropical disease American Economic Review Papers and Proceedings 99, 218-223.

Bleakley, H. 2007. Disease and development: Evidence from hookworm eradication in the american south Quarterly Journal of Economics 122, 73-117.

Boon, A.C., de Mutsert, G., van Baarle, D., Smith, D.J., Lapedes, A.S., Fouchier, R.A.M., Sintnicolaas, K., Osterahus, A.D., Rimmelzwaan, G.F. 2004. Recognition of homo- and heterosubtypic variants of influenza A viruses by human CD8+ T lymphocytes The Journal of Immunology 172, 2453-2460.

Burlando, A. 2009. Does malaria reduce education? findings from the ethiopian highlands.

Carrat, F., Lavenu, A., Epstein, S.L. 2006. Heterosubtypic immunity to influenza: Right hypothesis, wrong comparison [with reply to carrat and lavenu] The Journal of Infectious Diseases 193, 1613-1614.

Case, A., Fertig, A., Paxson, C. 2004. The lasting impact of childhood health and circumstance Journal of Health Economics 24, 365-389.

Centers for Disease Control and Prevention (CDC). 2011. Seasonal influenza: Questions & answers 2013.

Cox, N.J., Brammer, T.L., Postema, A., Kilmov, A.I., Thopmson, W., Fukuda, K. 2001. Influenza present: The impact of the 1999-2000 epidemic on morbidity and mortality in the USA International Congress Series 1219, 37-42.

Cox, N.J., Fukuda, K. 1998. Influenza Infectious Diseases of North America 12, 27-38.

Currie, J. 2011. Inequality at birth: Some causes and consequences American Economic Review Papers and Proceedings 101, 1-22.

Currie, J., Neidell, M. 2005. Air pollution and infant health: What can we learn from california's recent experience? Quarterly Journal of Economics 120, 1003-1030.

Currie, J., Neidell, M., Schmieder, J.F. 2009. Air pollution and infant health: Lessons from new jersey Journal of Health Economics 28, 688-703.

Currie, J., Stabile, M., Manivong, P., Roos, L.L. 2010. Child health and young adult outcomes Journal of Human Resources 45, 517-548.

Denoued, L., Turbelin, C., Ansart, S., Valleron, A., Flahault, A., Carrat, F. 2007. Predicting pneumonia and influenza mortality from morbidity data PLOS One 2, e464.

Domínguez, A., Muñoz, P., Martínez, A., Orcau, A. 1996. Monitoring mortality as an indicator of influenza in catalonia, spain Journal of Epidemiology and Community Health 50, 293-298.

Epstein, S.L. 2006. Prior H1N1 influenza infection and susceptibility of cleveland family study participants during the H2N2 pandemic of 1957: An experiment of nature The Journal of Infectious Diseases 193, 49-53.

Evans, A.S., Kaslow, R.A. Eds. 1997. Viral Infections of Humans, New York and London, .

Ferguson, N.M., Galvani, A.P., Bush, R.M. 2003. Ecological and immunological determinants of influena evolution Nature 422, 428-433.

Fiore, A.E., Uyekia, T.M., Broder, K., Finelli, L., Euler, G.L., Singleton, J.A., Iskander, J.H., Wortley, P.M., Shay, D.K., Bresee, J.S., Cox, N.J. 2010. Prevention and control of influenza with vaccines Morbidity and Mortality Weekly Report 59, 1-62.
Glezen, W.P., Couch, R.B. 1997. Influenza viruses, in: Evans, A.S., Kaslow, R.A. (Eds.), Viral Infections of Humans, New York and London, pp. 473-505.
Koelle, K., Cobey, S., Grenfell, B., Pascual, M. 2006. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans Science 314, 1898-1903.
Lavenu, A., Valleron, A., Carrat, F. 2004. Exploring cross-protection between influenza strains by an epidemiological model Virus Research 103, 101-105.
Lemaitre, M., Carrat, F. 2010. Comparative age distribution of influenza morbidity and mortality during seasonal influenza epidemics and the 2009 H1N1 pandemic BMC Infectious Disease 10, 162.
Maccini, S., Yang, D. 2009. Under the weather: Health, schooling, and economic consequences of early-life rainfall American Economic Review 99, 1006-1026.
Miguel, E., Kremer, M. 2004. Worms: Identifying impacts on education and health in the presence of treatment externalities Econometrica 72, 159-217.
Minnesota Population Center, State Health Access Data Assistance Center. 2012. Integrated Health Interview Series: Version 5.0, Minneapolis, MN, .
Molinari, N.M., Ortega-Sanchez, I.R., Messonnier, M.L., Thompson, W.W., Wortley, P.M., Weintraub, E., Bridges, C.B. 2007. The annual impact of seasonal influenza in the US: Measuring disease burden and costs Vaccine 25, 5086-5096.
Monto, A.S. 2008. Epidemiology of influenza Vaccine 26, D45-D48.
Oreopoulos, P., Page, M.E., Stevens, A.H. 2006. The intergenerational effects of compulsory schooling Journal of Labor Economics 24, 729-760.
Oreopoulos, P., Stabile, M., Walld, R., Roos, L.L. 2008. Short-, medium-, and long-term consequences of poor infant health: An analysis using siblings and twins Journal of Human Resources 43, 88-138.
Oswald, B.P., Ridenhour, B.J. 2013. Impact and cost-effectiveness of influenza vaccination in the united states 2005-2010 Working Paper.
Ruggles, S., Alexander, J.T., Genadek, K., Goeken, R., Schroeder, M.B., Sobek, M. 2010. Integrated Public use Microdata Series: Version 5.0 [Machine-Readable Database], Minneapolis, MN, .
Russel, C.A., Jones, T.C., Barr, I.G., Cox, N.J., Garten, R.J., Gregory, V., Gust, I.D., Hampson, A.W., Hay, A.J., Hurt, A.C., de Jong, J.C., Kelso, A., Klimov, A.I., Kageyama, T., Komadina, N., Lapedes, A.S., Lin, Y.P., Mosterin, A., Obuchi, M., Odagiri, T., Osterahus, A.D., Rimmelzwaan, G.F., Shaw, M.W., Skepner, E., Stohr, K., Tashiro, M., Fouchier, R.A.M., Smith, D.J. 2008. The global circulation of seasonal influenza A (H3N2) viruses Science 320, 340-346.
Seo, S.H., Webster, R.G. 2001. Cross-reactive, cell-mediated immunity and protection of chickens from lethal H5N1 influenza virus infection in hong kong poultry markets Journal of Virology 75, 2516-2525.
Simonsen, L. 2001. Influenza-related morbidity and mortality among children in developed and developing countries International Congress Series 1219, 13-19.
Smith, D.J. 2006. Predictability and preparedness in influenza control Science 312, 392-394.
Smith, D.J., Lapedes, A.S., de Jong, J.C., Bestebroer, T.M., Rimmelzwaan, G.F., Osterahus, A.D., Fouchier, R.A. 2004. Mapping the antigenic and genetic evolution of influenza virus Science 305, 371-376.
Sočan, M., Prosenc, K., Nagode, M. 2010. Differences in clinical predictors of influenza in adults and children with influenza-like illness Central European Journal of Medicine 5, 41-48.
Thompson, W.W., Shay, D.K., Weintraub, E., Brammer, L., Bridges, C.B., Cox, N.J., Fukuda, K. 2004. Influenza-associated hospitalizations in the united states Journal of the American Medical Association 292, 1333-1340.
Wilschut, J.C., McElhaney, J.E., Palache, A.M. 2006. Influenza, Edinburgh, U.K., .
Yetter, R.A., Barber, W.H., Small, P.A., Jr. 1980. Heterotypic immunity to influenza in ferrets Infection and Immunity 29, 650-653.

B. Multiple Strains, Attenuated Education Benefits, and Other Predictions

The exercise discussed in section 2 can be expanded to include multiple versions of influenza. The three most common type-strain combinations are A-H1N1, A-H3N2, and B. For the purposes of this extension, A-H1N1 and A-H3N2 are sufficiently similar such that experiencing one implies a high level of immunity, but not complete.[44] Experiencing one type-strain will provide complete immunity to that specific strain, and it provides partial immunity to strains with similar antigenic structure (Epstein, 2006, Smith et al., 2004, Smith, 2006, Russel et al., 2008, CDC, 2011, Fiore et al., 2010, Glezen, Couch, 1997). Symptoms experienced under A-H3N2 are the most severe for all populations (Monto, 2008, Simonsen, 2001). [45] The others are generally less severe with A-H1N1 still more severe than B, and this is still true across age groups (Wilschut et al., 2006). Assume that both A types are sufficiently different from B that there are minimal immunity gains from being exposed to one and experiencing the other. Consider equation (1) modified to account for multiple strains as (4):

$$(4)\ Education_i = Influenza_i^{prior\ exposure,strain} \times \theta^{prior\ exposure,strain}$$

where ^{strain} represents the dominant influenza strains during the period: A-H1N1, A-H3N2, and B. $^{prior\ exposure}$ represents previous exposure to influenza, for the sake of simplicity this will either be A-H1N1 or null.[46] $Influenza_i^{prior\ exposure,strain}$ is a 6 x A matrix of variables, and $\theta^{prior\ exposure,strain}$ is an A x 6 matrix of coefficients.[47]

There are three cases to consider, in which the exposure at age 11 will vary. In each case E will be exposed to influenza A-H1N1 at age four and U will remain unexposed. In the first case, at age 11 each is exposed to A-H1N1:[48]

$$\theta_{11}^{H1N1} < \theta_{11}^{H1N1,H1N1} = 0 = \theta_4^{H1N1} < \theta_4^{H1N1,H1N1}$$

This is the situation that we had previously.[49] U suffers from the exposure at age 11 (a negative relationship), meanwhile E is shielded by prior morbidity (a null effect). At age four, U is unaffected (a zero coefficient) and E benefits (a positive relationship). The next case to consider is exposure to A-H3N2 at age 11:

$$\theta_{11}^{H3N2} < \theta_{11}^{H1N1,H3N2} < \theta_4^{H3N2} = 0 < \theta_4^{H1N1,H3N2}$$

[44] Epstein (2006) found that adults previously exposed to H1N1 did not suffer any morbidity from H2N2 during the 1957 Asian Flu pandemic, despite high rates of exposure to H2N2, which was novel at the time.
[45] Increased morbidity in terms of both number of children infected and severity of symptoms experienced.
[46] In which case prior exposure will be blank and only the current strain presented.
[47] A for each strain previous immunity combination, or 3 for each age for E and 3 for each age for U.
[48] Where $^{H1N1,\ H1N1}$ represents exposure to A-H1N1 at both 4 and 11, respectively, for E, and where ^{H1N1} represents exposure to A-H1N1 at 11 only for U.
[49] Alternatively, we could think of this in terms of conditional expectations:

$$E[education|flu_4 = \emptyset \cap flu_{11} = H1N1] < E[education|flu_4 = H1N1 \cap flu_{11} = H1N1\ ] = 0 = E[educatiion|flu_4 = \emptyset] < E[education|flu_4 = H1N1]$$

U is in the same situation as before. E is now worse off, with there still being an advantage of early exposure, but it is lessened since E still experiences some morbidity at age 11 (a non-zero effect). This then implies that the age 4 coefficient will be smaller but still positive. Finally, consider the case were both are exposed to influenza B at age 11:

$$\theta_{11}^{B} \leq \theta_{11}^{H1N1,B} < \theta_{4}^{B} = 0 \leq \theta_{4}^{H1N1,B}$$

E has become less distinguishable from U.

The above three cases imply the following rankings for the coefficients (note that $\theta_a^B = \theta_a^{H1N1} = \theta_a^{H3N2}$):

$$\theta_{4}^{B} = 0 \leq \theta_{4}^{H1N1,B} < \theta_{4}^{H1N1,H3N2} < \theta_{4}^{H1N1,H1N1}$$

$$\theta_{11}^{B} \leq \theta_{11}^{H1N1,B} < \theta_{11}^{H1N1,H3N2} < 0 = \theta_{11}^{H1N1,H1N1}$$

where at age four parameters are (weakly) positive and at age eleven they are (weakly) negative. Early exposure to an influenza strain can be beneficial, however due to the continuum of immunity, the benefit is lessened and detrimental effects are still possible at older ages despite prior exposure.

The extension to the previous thought experiment presents a testable hypothesis: that individuals who are not previously exposed should suffer effects that are greater in magnitude than those that have been previously been exposed. Alternatively, one's first exposure to influenza will have a greater impact than subsequent exposures.

The alteration considered three different varieties of influenza that ranged in similarity. In reality there are a multitude of varieties even within a type-strain. As the virus becomes less detectable by the human immune system, immunization effects will move to zero and the negative effect of the shock itself will become more pronounced, i.e. the thought experiment of exposure to A-H1N1 at four and then influenza B at eleven becomes more relevant. The extent to which this can occur is unclear.

C. Cross-Type Influenza Immunity Exists

The acquisition of immunity being a channel through which educational attainment can benefit from exposure/morbidity only makes sense if: (a) influenza fails to vary significantly temporally, which is not the case (Smith, 2006, Smith et al., 2004, Russel et al., 2008).[50] (b) Or if immunity acquisition is more than strain-specific (i.e. cross-type or cross-subtype). There is

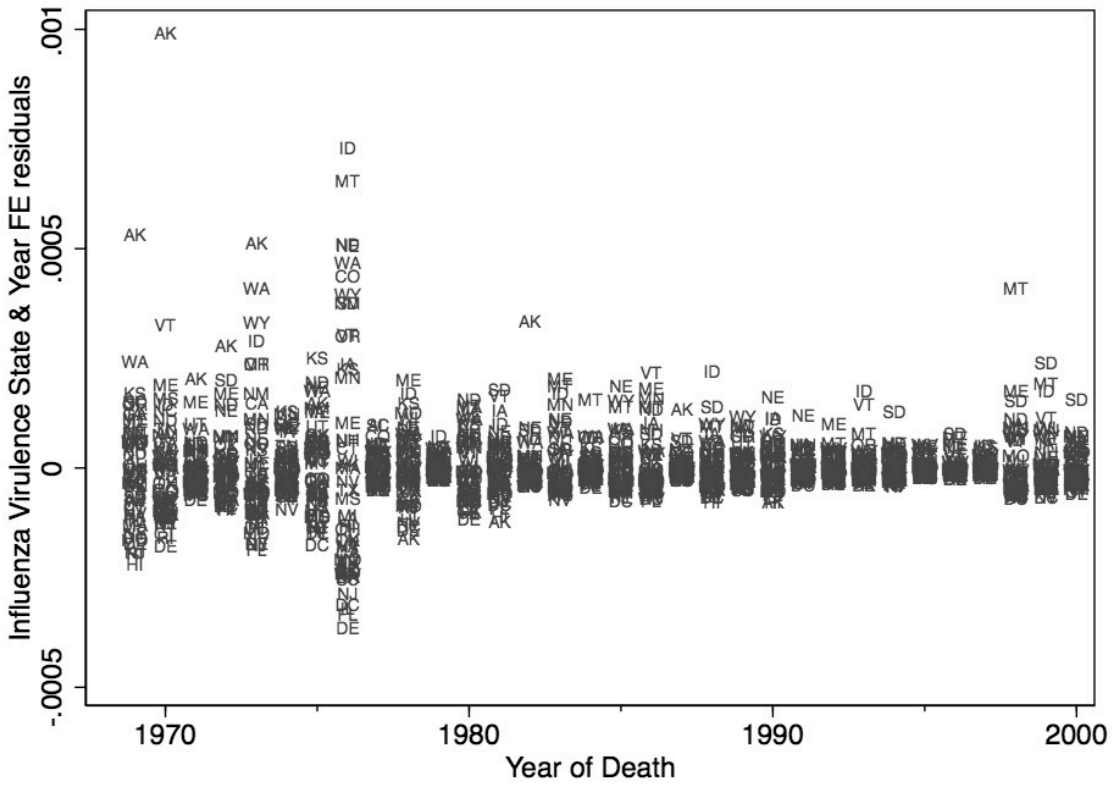
[50] For example, A-H3N2 changed sub-types every two to eight (mean of 3.4) years between 1968 and 2002 (Smith, 2006, Smith et al., 2004, Russel et al., 2008). The virus changed sufficiently to motivate the use of a new vaccine, on mean, twice in a three-year period (Smith et al., 2004).

ample evidence that this is so: in animals (Seo, Webster, 2001, Yetter et al., 1980, Glezen, Couch, 1997); in humans, via volunteer studies conducted in the mid-20th century, see Glezen & Couch (1997) for a comprehensive review; in humans, via clinical trials (Glezen, Couch, 1997, Epstein, 2006, Fiore et al., 2010); in humans, via laboratory experiments at the cellular level (Boon et al., 2004).[51] There is also theoretical/epidemiological support for cross-type influenza immunity in the explanation of otherwise anomalous phenomena (Carrat et al., 2006, Lavenu et al., 2004, Ferguson et al., 2003). Despite the preponderance of evidence in favor cross-type influenza immunity, this is still an active area of research and debate (Koelle et al., 2006, Russel et al., 2008).[52] Nonetheless, the CDC recognizes the existence and potential for cross-type immunity (CDC, 2011, Fiore et al., 2010).

[51] For human volunteer studies, the volunteers were infected with a given strain of influenza, then tested for immunity to strains. Clinical studies focused on known or determined prior morbidity, and subsequent morbidity based on exposure.
[52] The debate focuses on the need for more/better data, magnitude of immunity (rather than existence), and mechanism/pathway. The research now focuses on the efficacy of vaccines, rather than innate and morbidity-acquired immunity.

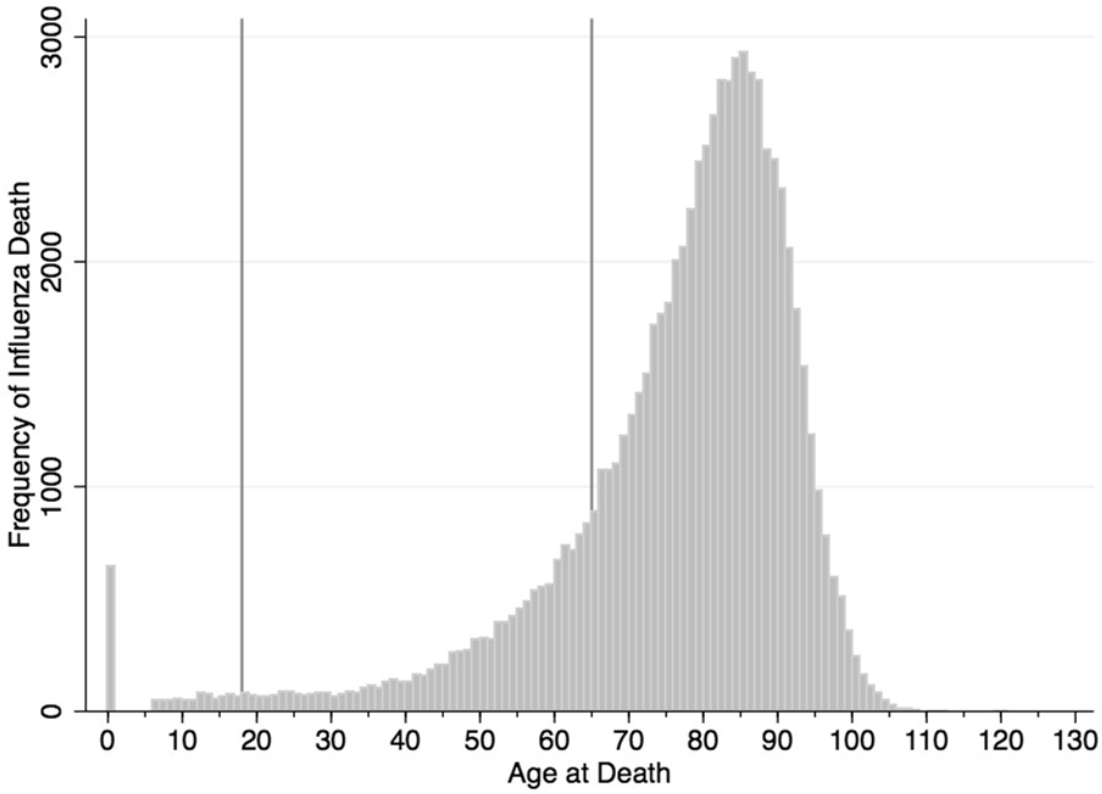
D. Figures

Figure 1 – Variation in Influenza Virulence Residuals, 1969-2000



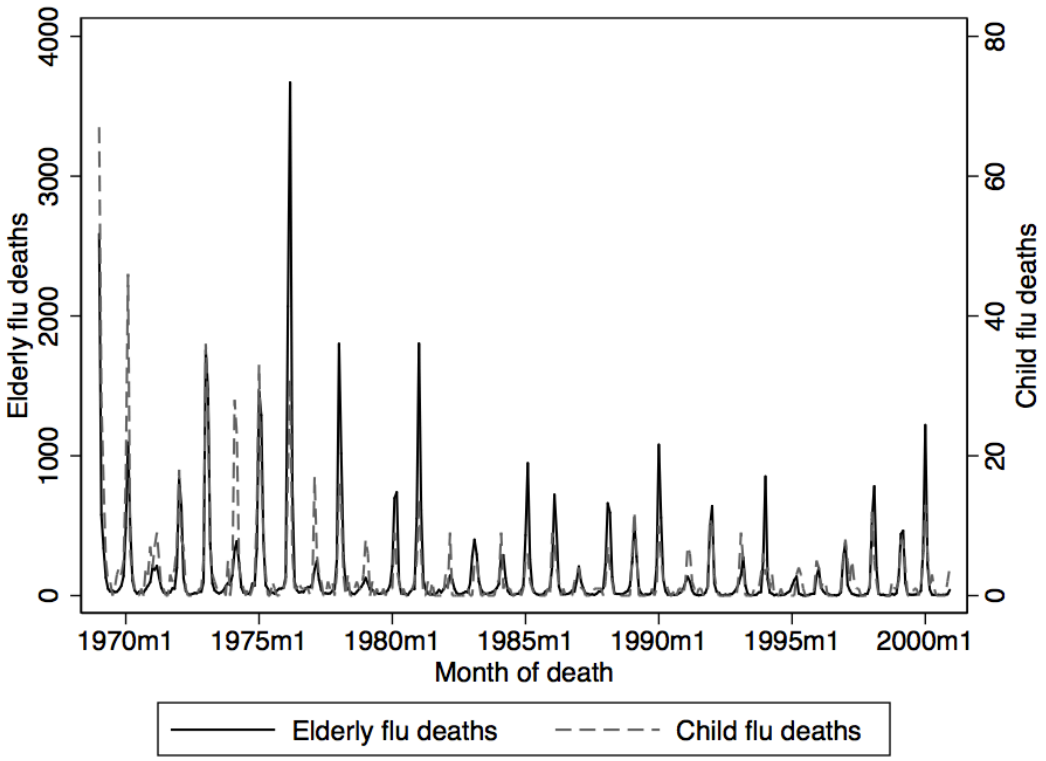
Notes: Influenza virulence is defined as elderly influenza deaths divided by the elderly population for a state and year. Flu deaths are taken from the Multiple Cause of Death (MCOD) and elderly population is taken from the Survey of Epidemiology and End Reports (SEER). Virulence is the elderly flu death risk for a given state and year of death. Residuals are presented with state and year fixed effects removed. The estimating equation is presented in footnote 11, see section 3 for more detail.

Figure 2 - Total Influenza Deaths by Age, 1968-2000



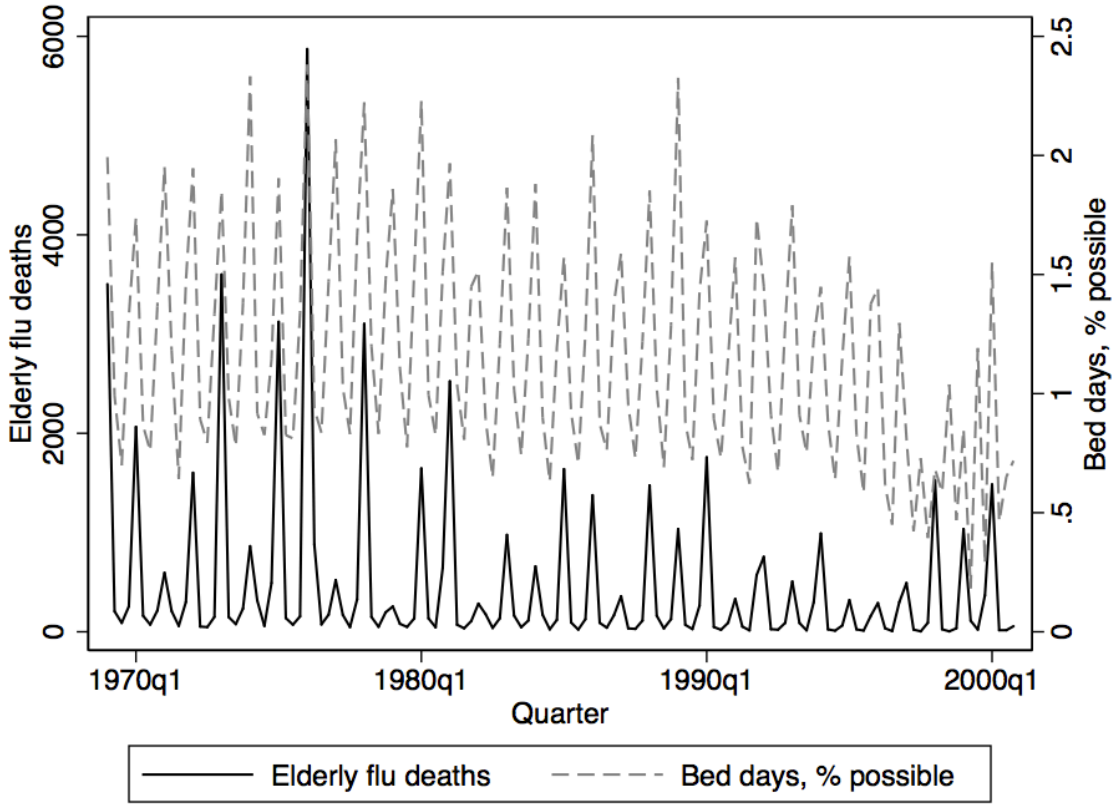
Notes: Influenza deaths are taken from the Multiple Cause of Death (MCOD). Vertical lines at 18 and 65 represent the end of childhood and the beginning of seniorhood, respectively. It is the elderly who are most likely to die from influenza. After infancy, children do not die from the virus. Between 1968 and 2000, there were only 1,256 child deaths attributable to influenza. During this time there were approximately 1 million child deaths.

Figure 3 – Elderly & Child Influenza Monthly Mortality, 1969-2000



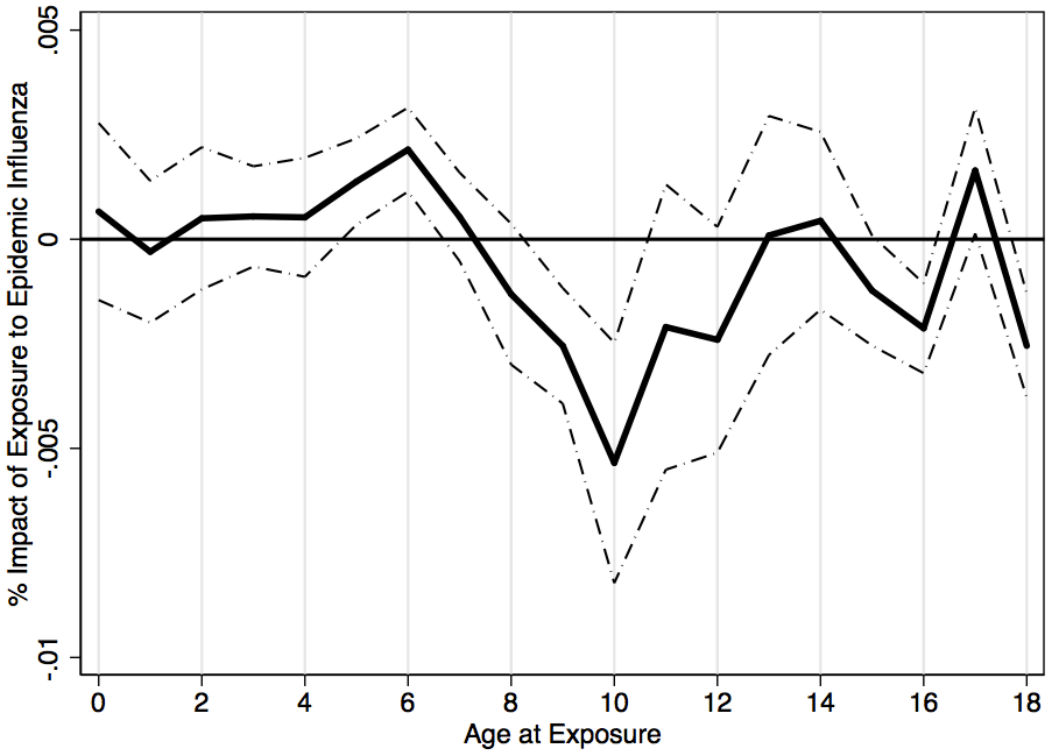
Notes: Elderly and child flu deaths are taken from the Multiple Cause of Death (MCOD). These two trends have a correlation of 0.7915. The flu season is October to April; it peaks in January. The flu season is the same for the elderly and children.

Figure 4 – Elderly Mortality and Child Morbidity (Days Sick in Bed), 1969-2000



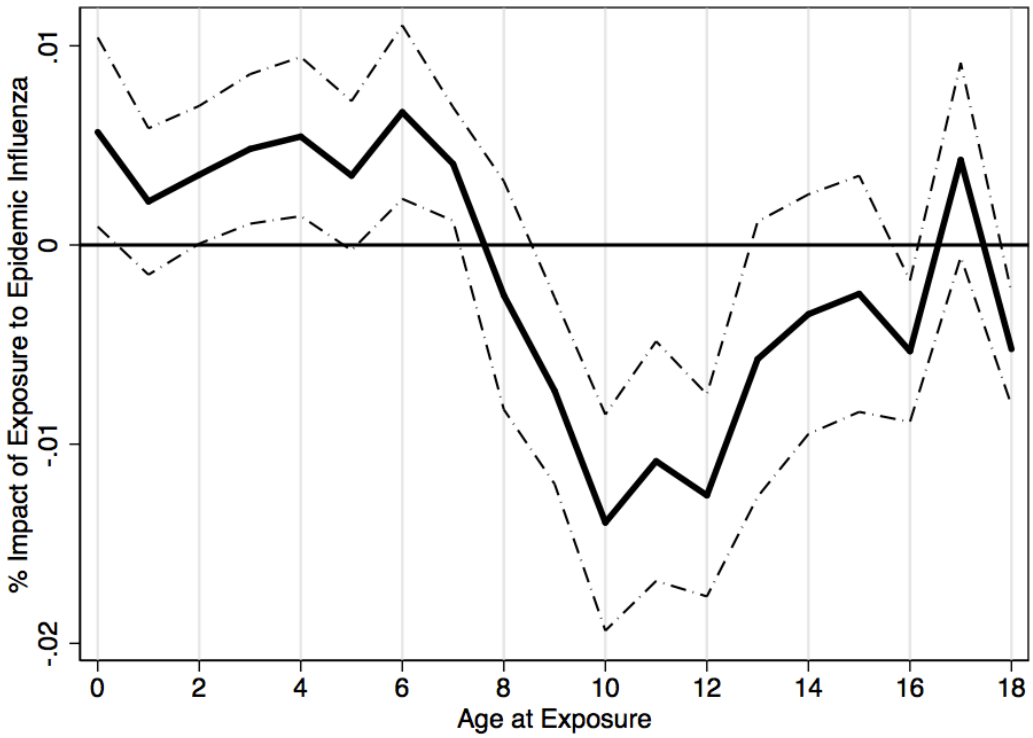
Notes: Elderly flu deaths are taken from the Multiple Cause of Death (MCOD). Bed days were obtained from the National Health Interview Survey (NHIS), which is a quarterly survey. Bed days measure the number of days that a child spends sick in bed for any illness. Prior to 1997 bed days referred to the previous two weeks; beginning in 1997 they regard the previous year. Hence, using a percentage measure. The correlation between these two quarterly series is 0.6343

Figure 5 – Impact of Influenza on Obtaining a High School Diploma (or equivalent)



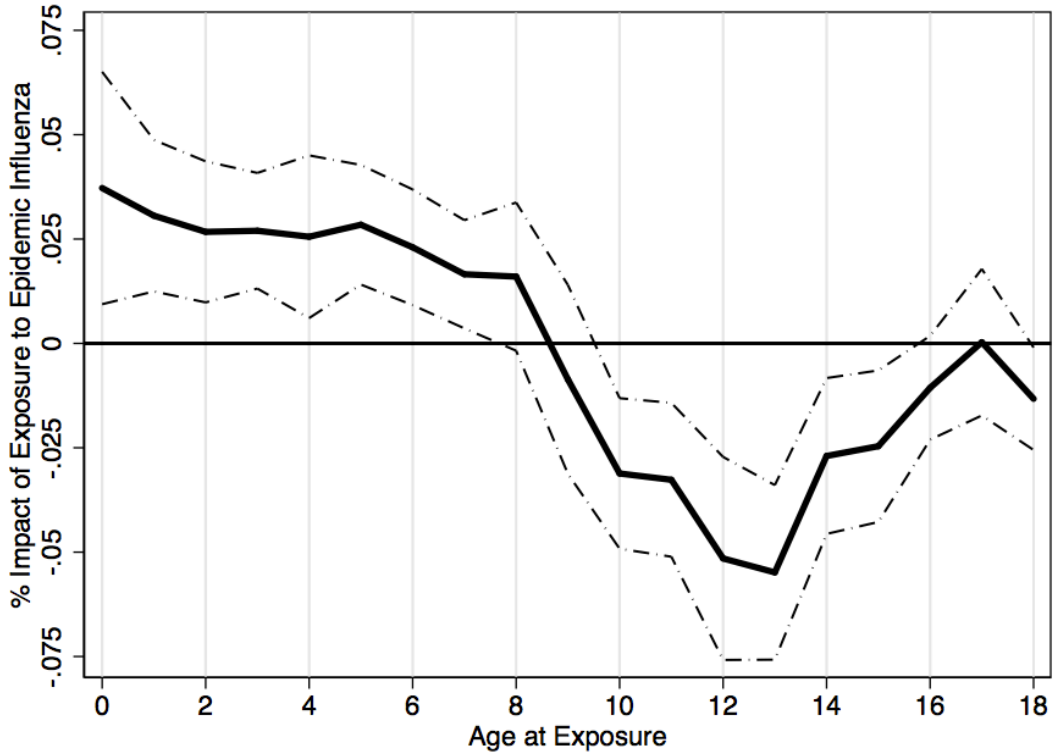
Notes: Epidemic influenza is defined using the interquartile range of influenza virulence. The 95% confidence interval is shown in dot-dash. See section 5 for greater detail.

Figure 6 – Impact on Influenza on Obtaining Some College



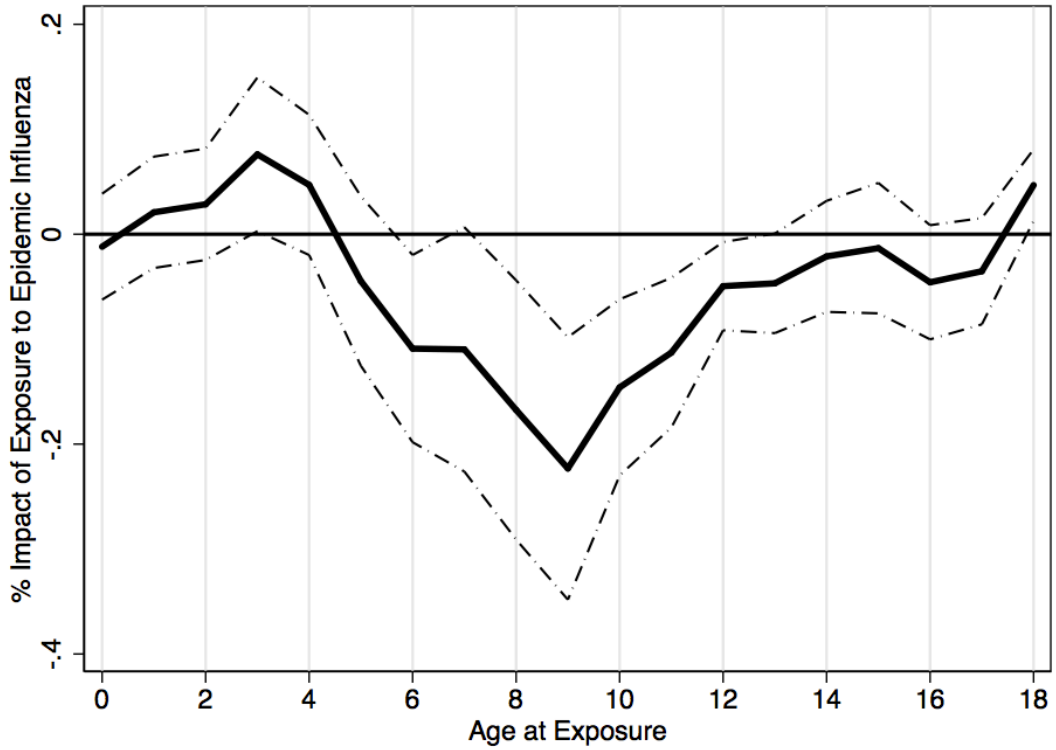
Notes: Epidemic influenza is defined using the interquartile range of influenza virulence. The 95% confidence interval is shown in dot-dash. See section 5 for greater detail.

Figure 7 – Impact of Influenza on Obtaining a Bachelor's Degree



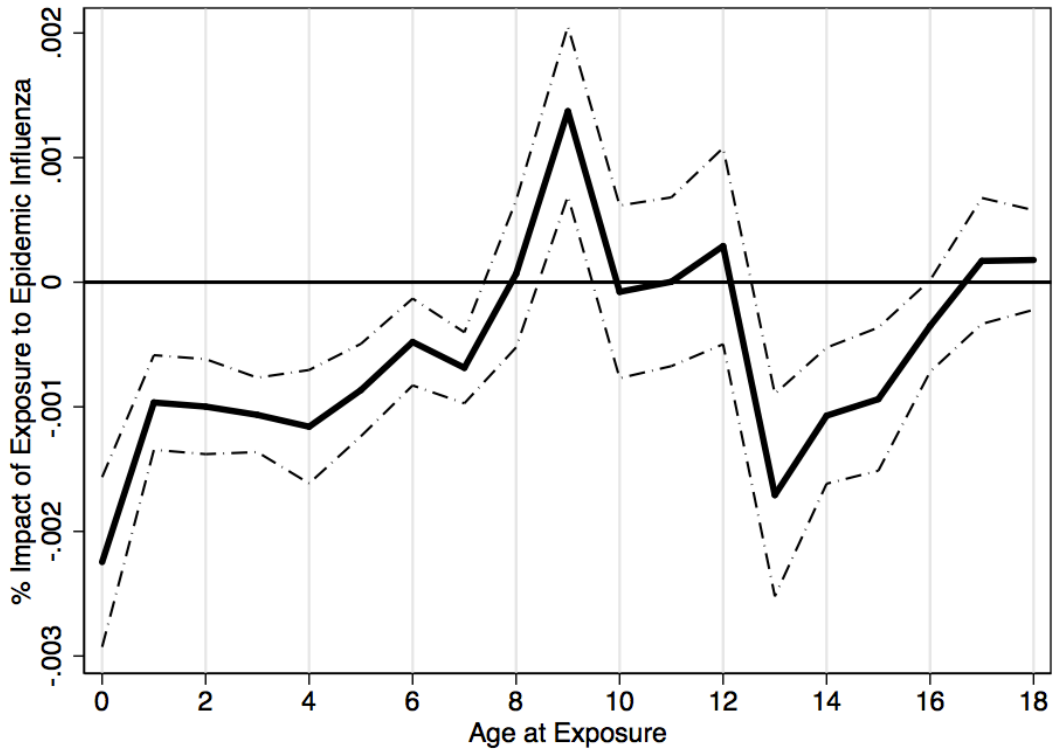
Notes: Epidemic influenza is defined using the interquartile range of influenza virulence. The 95% confidence interval is shown in dot-dash. See section 5 for greater detail.

Figure 8 – Impact of Influenza on Attending College



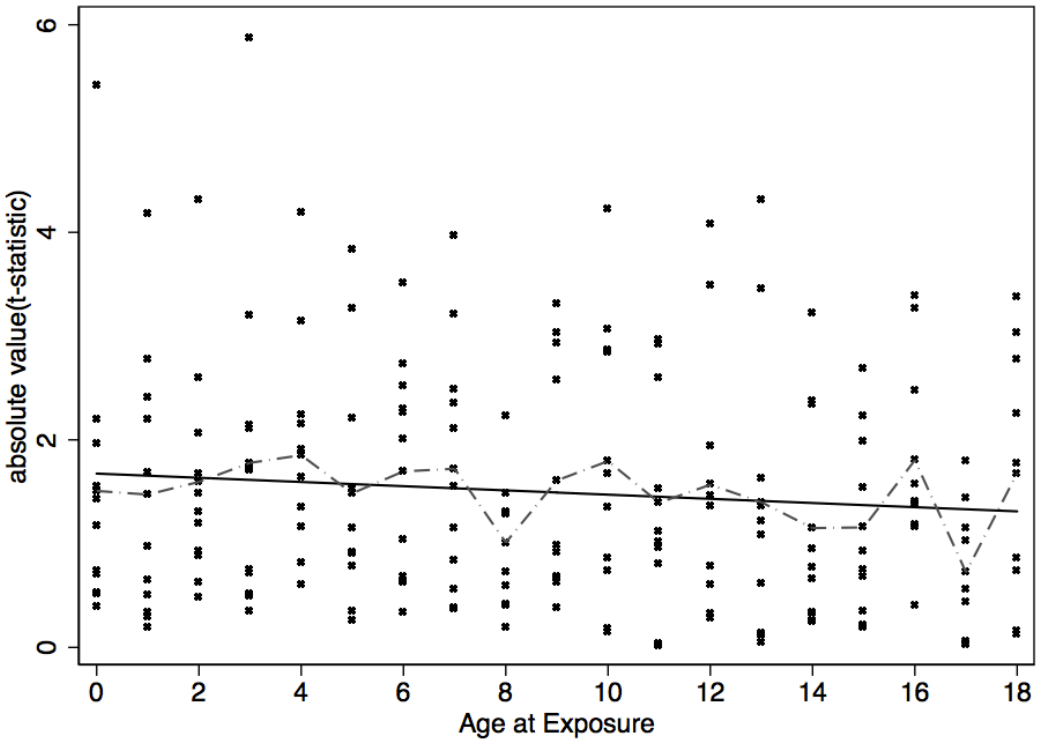
Notes: Epidemic influenza is defined using the interquartile range of influenza virulence. The 95% confidence interval is shown in dot-dash. See section 5 for greater detail.

Figure 9 – Impact of Influenza on the Absence of Cognitive Difficulty



Notes: Cognitive difficulties are presented as a bad in the Census. To instead represent a good, the results have been multiplied by negative one. This follows the same form as the educational outcomes (a positive coefficient increases a good). Epidemic influenza is defined using the interquartile range of influenza virulence. The 95% confidence interval is shown in dot-dash. See section 5 for greater detail.

Figure 10 – T-Statistics Against Age at Exposure



Notes: X's represent the absolute value of t-statistis from estimating (2), as taken from Tables 2–4. The solid-black line is a univariate linear regression of age at exposure on the absolute value of the t-statistic (see footnote 39 for the estimating equation). The slope coefficient is -0.0298 with a t-statistic of -8.80. The gray-dash-dot line is the mean of the absolute value of the t-statistics at a given age.

E. Tables

Table 1 – Summary Statistics

	All		Non-Movers		Movers	
	Mean	Stan. Dev.	Mean	Stan. Dev.	Mean	Stan. Dev.
Outcomes:						
1(on-time grade, ages 7–11) *	0.8168	0.3869	0.8182	0.3857	0.8108	0.3917
1(on-time grade, ages 15–19)*	0.7439	0.4265	0.7434	0.4260	0.7420	0.4288
1(high school)	0.8751	0.3307	0.8607	0.3463	0.9049	0.2934
1(some college)	0.5911	0.4916	0.5506	0.4974	0.6752	0.4683
1(bachelor's)	0.2485	0.4321	0.2076	0.4056	0.3334	0.4714
1(attending college)*	0.1026	0.3035	0.1015	0.3020	0.1050	0.3065
1(attending graduate school)	0.0379	0.1910	0.0301	0.1710	0.0541	0.2263
1(disability working, 2000)	0.0379	0.1910	0.0301	0.1710	0.0541	0.2263
1(cognitive difficulty)	0.0308	0.1727	0.0320	0.1761	0.0281	0.1654
1(ambulatory difficulty)	0.0262	0.1597	0.0276	0.1638	0.0233	0.1508
1(seeing/hearing difficulty)	0.0145	0.1195	0.0147	0.1202	0.0141	0.1181
Influenza Exposure at:						
Infancy	0.00072	0.00093	0.00074	0.00094	0.00067	0.00091
1 year old	-	-	0.00063	0.00088	-	-
2 years old	-	-	0.00064	0.00089	-	-
3 years old	-	-	0.00062	0.00089	-	-
4 years old	-	-	0.00063	0.00091	-	-
5 years old	-	-	0.00058	0.00086	-	-
6 years old	-	-	0.00056	0.00090	-	-
7 years old	-	-	0.00049	0.00087	-	-
8 years old	-	-	0.00031	0.00042	-	-
9 years old	-	-	0.00032	0.00043	-	-
10 years old	-	-	0.00026	0.00031	-	-
11 years old	-	-	0.00026	0.00032	-	-
12 years old	-	-	0.00024	0.00029	-	-
13 years old	-	-	0.00021	0.00023	-	-
14 years old	-	-	0.00024	0.00027	-	-
15 years old	-	-	0.00023	0.00027	-	-
16 years old	-	-	0.00022	0.00026	-	-
17 years old	-	-	0.00020	0.00024	-	-
18 years old	-	-	0.00019	0.00023	-	-
Total Childhood Influenza Exposure	-	-	0.00759	0.00691	-	-
Controls:						
Age	26.64	2.32	26.58	2.32	26.76	2.31
Birth Year	1972.4	2.3	1972.4	2.3	1972.2	2.3
1(white)	0.7844	0.4112	0.7710	0.4202	0.8123	0.3904
1(black)	0.1323	0.3388	0.1442	0.3513	0.1075	0.3097
1(other)	0.0722	0.2588	0.0735	0.2610	0.0693	0.2540
1(male)	0.4942	0.5000	0.4937	0.5000	0.4952	0.5000
Birth State == Residence	0.6749	0.4684	1.0000	0.0000	0.0000	0.0000
Observations	1,072,180		723,636		348,544	
Sample			$23 \leq \text{Age} \leq 31$ $1969 \leq \text{Birth Year} \leq 1977$			

Notes: The samples for 1(on-time grade, ages 7–11), 1(on-time grade, ages 15–19), and 1(attending college) are different than those listed for the remainder of Table 1. The specifics of the samples are available in section 4, and all following tables. Note that

unless the state of birth and residence are identical it is not meaningful to construct a history of childhood influenza exposure without additional assumptions hence it is omitted here. The birth state is known, so exposure at infancy can be meaningfully presented. This data comes from the 2000 U.S. Census, except as noted previously. Influenza exposure is constructed using elderly influenza deaths using the Multiple Cause of Death (MCOD).

Table 2 – Complete Childhood Influenza Exposure on Educational Attainment

	Educational Attainment		
Influenza exposure at:	1(high school)	1(some coll)	1(bachelors)
Infancy	0.00066	0.00566**	0.03722**
	(0.6069)	(0.0495)	(0.0279)
1 year old	-0.00029	0.00218	0.03059***
	(0.7714)	(0.3287)	(0.0055)
2 years old	0.00050	0.00352*	0.02671***
	(0.6285)	(0.0938)	(0.0093)
3 years old	0.00054	0.00481**	0.02698***
	(0.4543)	(0.0347)	(0.0014)
4 years old	0.00052	0.00544**	0.02555**
	(0.5446)	(0.0250)	(0.0310)
5 years old	0.00138**	0.00347	0.02842***
	(0.0276)	(0.1267)	(0.0011)
6 years old	0.00214***	0.00666**	0.02302***
	(0.0004)	(0.0118)	(0.0063)
7 years old	0.00053	0.00406**	0.01656**
	(0.4050)	(0.0189)	(0.0355)
8 years old	-0.00131	-0.00253	0.01600
	(0.1994)	(0.4678)	(0.1381)
9 years old	-0.00254***	-0.00731**	-0.00861
	(0.0024)	(0.0101)	(0.5311)
10 years old	-0.00534***	-0.01393***	-0.03116***
	(0.0022)	(0.0000)	(0.0045)
11 years old	-0.00209	-0.01085***	-0.03266***
	(0.3113)	(0.0030)	(0.0035)
12 years old	-0.00240	-0.01257***	-0.05151***
	(0.1431)	(0.0000)	(0.0005)
13 years old	0.00009	-0.00572	-0.05485***
	(0.9587)	(0.1723)	(0.0000)
14 years old	0.00044	-0.00347	-0.02696**
	(0.7324)	(0.3420)	(0.0175)
15 years old	-0.00122	-0.00245	-0.02462**
	(0.1246)	(0.4963)	(0.0257)
16 years old	-0.00212***	-0.00533**	-0.01063
	(0.0011)	(0.0134)	(0.1593)
17 years old	0.00164*	0.00427	0.00028
	(0.0726)	(0.1496)	(0.9787)
18 years old	-0.00254***	-0.00522***	-0.01325*
	(0.0007)	(0.0024)	(0.0757)
Controls:			
1($state_s$)	X	X	X
1($race_r$)	X	X	X
1(male)	X	X	X
1(birth $year_{yr}$)	X	X	X
Observations:	723,636	723,636	723,636
Sample:			
Census(es)	2000	2000	2000
Ages	23 – 31	23 – 31	23 – 31
Birth Years	1969 – 1977	1969 – 1977	1969 – 1977

Notes: Adjusted coefficients are presented, see section 5, providing the percent change due to exposure to epidemic levels of influenza. p-values from state clustered standard errors are presented in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 3 – Complete Childhood Influenza Exposure on Education Mechanisms

	Grade-for-Age		Attending School	
Influenza exposure at:	7–11	15–19	1(college)	1(graduate)
Infancy	-0.00441	0.04328	-0.01190	0.06118
	(0.6021)	(0.1209)	(0.6979)	(0.1543)
1 year old	-0.01377**	-0.00400	0.02083	0.06782**
	(0.0279)	(0.8456)	(0.5186)	(0.0160)
2 years old	-0.00351	0.01818	0.02848	0.03891
	(0.5308)	(0.3541)	(0.3772)	(0.1381)
3 years old	-0.00434	-0.02644**	0.07605*	0.05473*
	(0.6207)	(0.0326)	(0.0875)	(0.0851)
4 years old	-0.01232	0.02549***	0.04696	0.04144*
	(0.1753)	(0.0017)	(0.2477)	(0.0560)
5 years old	-0.01012	-0.01154	-0.04467	0.02203
	(0.2519)	(0.3604)	(0.3630)	(0.2502)
6 years old	-0.01076	-0.03316**	-0.10901**	0.01445
	(0.2968)	(0.0219)	(0.0445)	(0.4954)
7 years old	-0.00499	0.06738***	-0.10975	0.00709
	(0.2491)	(0.0013)	(0.1208)	(0.7149)
8 years old	-	0.00898	-0.16748**	0.04451
	-	(0.6906)	(0.0257)	(0.1929)
9 years old	-	-0.01015	-0.22321***	-0.04995
	-	(0.4967)	(0.0033)	(0.3618)
10 years old	-	0.04078*	-0.14600***	-0.03289
	-	(0.0934)	(0.0042)	(0.4606)
11 years old	-	0.00117	-0.11284***	-0.03981
	-	(0.9733)	(0.0093)	(0.3356)
12 years old	-	-0.04412	-0.04950*	-0.00947
	-	(0.1728)	(0.0529)	(0.7478)
13 years old	-	0.00317	-0.04682	-0.05701
	-	(0.9077)	(0.1044)	(0.2228)
14 years old	-	-0.00362	-0.02114	0.01907
	-	(0.7921)	(0.5110)	(0.7471)
15 years old	-	-0.02352	-0.01328	-0.01173
	-	(0.3524)	(0.7250)	(0.8440)
16 years old	-	-	-0.04582	-0.04498
	-	-	(0.1658)	(0.2460)
17 years old	-	-	-0.03528	-0.00142
	-	-	(0.2519)	(0.9744)
18 years old	-	-	0.04674**	-0.02369
	-	-	(0.0240)	(0.3869)
Controls:				
1($state_s$)	X	X	X	X
1($race_r$)	X	X	X	X
1(male)	X	X	X	X
1(birth $year_{yr}$)	X	X	X	X
1(birth $quarter_q$)	X			
Observations:	533,734	1,214,542	808,490	723,636
Sample:				
Census(es)	1980	1990 & 2000	2000	2000
Ages	7 – 11	15 – 19	18 – 25	23 – 31
Birth Years	1969 – 1973	1970 – 1974 & 1980 – 1984	1974 – 1981	1969 – 1977

Notes: Adjusted coefficients are presented, see section 5. p-values from state clustered standard errors are presented in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 4 – Complete Childhood Influenza Exposure on Adult Health

	Health			
Influenza exposure at:	1(diswrk00)	1(cog diff)	1(amb diff)	1(see/hear diff)
Infancy	0.00623	0.00224***	0.00034	0.00024
	(0.4838)	(0.0000)	(0.2423)	(0.4632)
1 year old	0.00986*	0.00096***	0.00012	-0.00007
	(0.0910)	(0.0000)	(0.6116)	(0.7395)
2 years old	0.01284**	0.00099***	0.00027	-0.00033
	(0.0396)	(0.0000)	(0.2330)	(0.1926)
3 years old	0.00306	0.00106***	-0.00006	0.00022
	(0.6096)	(0.0000)	(0.7298)	(0.4748)
4 years old	0.00744	0.00115***	0.00036	0.00014
	(0.2476)	(0.0000)	(0.1005)	(0.4137)
5 years old	0.00125	0.00086***	0.00012	-0.00006
	(0.7935)	(0.0001)	(0.4369)	(0.7290)
6 years old	0.00299	0.00047**	-0.00005	-0.00012
	(0.5291)	(0.0237)	(0.7352)	(0.5244)
7 years old	-0.00151	0.00068***	-0.00007	0.00025**
	(0.7062)	(0.0001)	(0.5753)	(0.0131)
8 years old	0.00397	-0.00006	0.00070	0.00026
	(0.6746)	(0.8514)	(0.1385)	(0.5521)
9 years old	-0.00209	-0.00137***	0.00041	-0.00016
	(0.7048)	(0.0009)	(0.3271)	(0.5098)
10 years old	0.00984	0.00007	0.00007	0.00050
	(0.3909)	(0.8543)	(0.8789)	(0.1777)
11 years old	0.00838	-0.00000	-0.00048	-0.00046
	(0.4222)	(0.9930)	(0.2658)	(0.1275)
12 years old	0.00261	-0.00028	-0.00056	0.00026
	(0.7776)	(0.5450)	(0.1746)	(0.4327)
13 years old	0.00641	0.00170***	0.00005	0.00031
	(0.5388)	(0.0005)	(0.8924)	(0.2767)
14 years old	-0.00618	0.00107***	0.00011	0.00075**
	(0.4400)	(0.0013)	(0.8029)	(0.0194)
15 years old	0.00174	0.00093***	0.00073**	0.00025
	(0.8317)	(0.0072)	(0.0469)	(0.4526)
16 years old	-0.01550***	0.00035	0.00012	0.00022
	(0.0007)	(0.1158)	(0.6852)	(0.2375)
17 years old	-0.00537	-0.00017	0.00002	0.00009
	(0.3065)	(0.5781)	(0.9535)	(0.6633)
18 years old	-0.00100	-0.00017	-0.00005	0.00064***
	(0.8716)	(0.4634)	(0.9019)	(0.0056)
Controls used:				
1($state_s$)	X	X	X	X
1($race_r$)	X	X	X	X
1(male)	X	X	X	X
1(birth $year_{yr}$)	X	X	X	X
Observations:	723,636	723,636	723,636	723,636
Sample:				
Census(es)	2000	2000	2000	2000
Ages	23 – 31	23 – 31	23 – 31	23 – 31
Birth Years	1969 – 1977	1969 – 1977	1969 – 1977	1969 – 1977

Notes: Adjusted coefficients are presented, see section 5. p-values from state clustered standard errors are presented in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 5 – Relationship Between Population Morbidity Rates and Children's Influenza

Flu Season is	% Population Ill with Flu	% Children Ill with Flu	Times Sick in Childhood
Mild	5.0%	9.3%	1.8
Median	12.5%	23.2%	4.4
Epidemic	20.0%	37.2%	7.1

Notes: % population ill with flu, η, is taken from Cox & Fukuda (1998) and Monto (2008). % children ill with flu uses η and Lemaitre & Carrat's (2010) relative illness ratio (RIR). Times sick in childhood is the sum of the probabilities of getting ill at each year during childhood. Since these are all mechanically linked, they can be used as alternative interpretations for the following tables. See section 5.2 for greater detail.

Table 6 – Aggregated Effect of Influenza on Educational Attainment

Influenza Virulence Throughout Childhood	Educational Attainment		
	1(high school)	1(some coll)	1(bachelors)
Mild	-0.00129	-0.00390	-0.00739
	(0.0000)	(0.0000)	(0.0000)
Median	-0.00322	-0.00975	-0.01848
	(0.0000)	(0.0000)	(0.0003)
Epidemic	-0.00515	-0.01560	-0.02956
	(0.0000)	(0.0000)	(0.0007)

Notes: Author calculated standard errors are presented in parenthesis. Significance stars are omitted, as all aggregate effects are highly significant. See section 5.2 for greater detail.

Table 7 – Aggregated Effect of Influenza on Educational Mechanisms

Influenza Virulence Throughout Childhood	Grade-for-Age		Attending School	
	7–11	15–19	1(college)	1(graduate)
Mild	-0.00541	0.00392	-0.09896	0.00291
	(0.0000)	(0.0000)	(0.0017)	(0.0006)
Median	-0.01351	0.00979	-0.24739	0.00728
	(0.0000)	(0.0000)	(0.0105)	(0.0040)
Epidemic	-0.02162	0.01567	-0.39583	0.01165
	(0.0001)	(0.0000)	(0.0268)	(0.0103)

Notes: Author calculated standard errors are presented in parenthesis. Significance stars are omitted, as all aggregate effects are highly significant. See section 5.2 for greater detail.

Table 8 – Aggregated Effect of Influenza on Adult Health

Influenza Virulence Throughout Childhood	Health			
	1(diswrk00)	1(cog diff)	1(amb diff)	1(see/hear diff)
Mild	-0.00406	-0.00088	-0.00018	-0.00028
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Median	-0.01015	-0.00220	-0.00045	-0.00070
	(0.0002)	(0.0000)	(0.0000)	(0.0000)
Epidemic	-0.01624	-0.00353	-0.00073	-0.00113
	(0.0005)	(0.0000)	(0.0000)	(0.0000)

Notes: Author calculated standard errors are presented in parenthesis. Significance stars are omitted, as all aggregate effects are highly significant. See section 5.2 for greater detail.

Table 9 – Birth Year Influenza Exposure on Educational Attainment

	Educational Attainment		
Influenza exposure at:	1(high school)	1(some coll)	1(bachelors)
Infancy	0.635	0.596	0.557
	(0.430)	(1.396)	(1.014)
Mean influenza	0.00072	0.00072	0.00072
Controls used:			
1($state_s$)	X	X	X
1($race_r$)	X	X	X
1(male)	X	X	X
1(birth $year_{yr}$)	X	X	X
Observations:	1,072,180	1,072,180	1,072,180
Sample:			
Census(es)	2000	2000	2000
Ages	23 – 31	23 – 31	23 – 31
Birth Years	1969 – 1977	1969 – 1977	1969 – 1977

Notes: State clustered standard errors are presented in parenthesis. *** p<0.01, **, p<0.05, * p<0.1

Table 10 – Birth Year Influenza Exposure on Educational Mechanisms

	Grade-for-Age		Attending School	
Influenza exposure at:	7–11	15–19	1(college)	1(graduate)
Infancy	0.427	-1.536	-2.322	0.821*
	(6.729)	(48.03)	(6.800)	(0.452)
Mean influenza	0.00058	0.00038	0.00072	0.00072
Controls used:				
1(state$_s$)	X	X	X	X
1(race$_r$)	X	X	X	X
1(male)	X	X	X	X
1(birth year$_{yr}$)	X	X	X	X
1(birth quarter$_q$)	X			
Observations:	659,267	1,573,794	1,124,555	1,072,180
Sample:				
Census(es)	1980	1990 & 2000	2000	2000
Ages	7 – 11	15 – 19	18 – 25	23 – 31
Birth Years	1969 – 1973	1970 – 1974 & 1980 – 1984	1974 – 1981	1969 – 1977

Notes: State clustered standard errors are presented in parenthesis. *** p<0.01, **, p<0.05, * p<0.1

Table 11 – Birth Year Influenza Exposure on Adult Health

	Health			
Influenza exposure at:	1(diswrk00)	1(cog diff)	1(amb diff)	1(see/hear diff)
Infancy	0.483*	-0.543**	0.236	-0.440
	(0.285)	(0.242)	(0.215)	(0.350)
Mean influenza	0.00072	0.00072	0.00072	0.00072
Controls used:				
1(state$_s$)	X	X	X	X
1(race$_r$)	X	X	X	X
1(male)	X	X	X	X
1(birth year$_{yr}$)	X	X	X	X
Observations:	1,072,180	1,072,180	1,072,180	1,072,180
Sample:				
Census(es)	2000	2000	2000	2000
Ages	23 – 31	23 – 31	23 – 31	23 – 31
Birth Years	1969 – 1977	1969 – 1977	1969 – 1977	1969 – 1977

Notes: State clustered standard errors are presented in parenthesis. *** p<0.01, **, p<0.05, * p<0.1

Table 12 – Comparing First Shock vs Any Shock for Total Effect for Education

	Median Influenza	
	Any Shock	First Shock
Educational Attainment		
1(high school)		
Infancy – 18 years	-0.0106	-0.0134
	(0.0001)	(0.0004)
Infancy – 8 years	-0.0028	-0.0134
	(0.0000)	(0.0004)
1(some college)		
Infancy – 18 years	-0.0123	-0.0070
	(0.0001)	(0.0002)
Infancy – 8 years	-0.0104	-0.0070
	(0.0000)	(0.0002)
1(bachelors)		
Infancy – 18 years	-0.0938	-0.1944
	(0.0002)	(0.0004)
Infancy – 8 years	-0.0650	-0.1944
	(0.0000)	(0.0004)

Notes: This tests the hypothesis that the first shock matters most, presented in appendix section B. This hypothesis is confirmed in 85% of the individual age coefficients. For ease of exhibition, only the aggregate impacts are presented here for median influenza levels during the entirety of childhood. Author calculated standard errors are presented in parenthesis.

Table 13 – Comparing First Shock vs Any Shock for Total Effect for Education Mechanisms

	Median Influenza	
	Any Shock	First Shock
Attending School		
1(college attend)		
Infancy – 18 years	-0.0415	0.5731
	-(0.0002)	-(0.0006)
Infancy – 8 years	-0.0683	0.5731
	(0.0000)	-(0.0006)
1(grad school attend)		
Infancy – 18 years	-0.0290	-0.1175
	(0.0000)	(0.0000)
Infancy – 8 years	-0.0134	-0.1175
	(0.0000)	(0.0000)
On-Time Grade Progress		
1(GfA, 7–11)		
Infancy – 18 years	0.0068	-0.0006
	(0.0000)	(0.0000)
Infancy – 8 years	0.0068	-0.0006
	(0.0000)	(0.0000)
1(GfA, 15–19)		
Infancy – 18 years	-0.0122	-0.0705
	-(0.0011)	-(0.0032)
Infancy – 8 years	-0.0115	-0.0705
	-(0.0003)	-(0.0032)

Notes: This tests the hypothesis that the first shock matters most, presented in appendix section B. This hypothesis is confirmed in 85% of the individual age coefficients. For ease of exhibition, only the aggregate impacts are presented here for median influenza levels during the entirety of childhood. Author calculated standard errors are presented in parenthesis.

Table 14 – Comparing First Shock vs Any Shock for Total Effect for Adult Health

	Median Influenza	
	Any Shock	First Shock
Health		
1(diswrk00)		
Infancy – 18 years	-0.0057	0.1880
	(0.0000)	(0.0000)
Infancy – 8 years	0.0006	0.1880
	(0.0000)	(0.0000)
1(cog diff)		
Infancy – 18 years	-0.0214	0.1348
	(0.0000)	(0.0000)
Infancy – 8 years	-0.0171	0.1348
	(0.0000)	(0.0000)
1(amb diff)		
Infancy – 18 years	0.0905	0.1893
	(0.0000)	(0.0000)
Infancy – 8 years	0.0348	0.1893
	(0.0000)	(0.0000)
1(see/hear diff)		
Infancy – 18 years	0.0315	0.1932
	(0.0000)	(0.0000)
Infancy – 8 years	0.0317	0.1932
	(0.0000)	(0.0000)

Notes: This tests the hypothesis that the first shock matters most, presented in appendix section B. This hypothesis is confirmed in 85% of the individual age coefficients. For ease of exhibition, only the aggregate impacts are presented here for median influenza levels during the entirety of childhood. Author calculated standard errors are presented in parenthesis.