Long-term income effects from early-life exposure to the 1918/1919 influenza pandemic: the case of southern Sweden

Tommy Bengtsson and Jonas Helgertz
Centre for Economic Demography and Department of Economic History
School of Economics and Management, Lund University

Abstract

The influenza pandemic of 1918/1919 (the "Spanish influenza") has been used as a quasi-experiment by Almond (2006) to test the *fetal origins hypothesis* on health and socioeconomic outcomes later in life, using samples from the 1960-1980 censuses for the U.S. The findings suggested that prenatal exposure to the influenza pandemic had negative effects on education, health and income attainment during adulthood (Almond 2006:708). The implications of these findings have recently been questioned by Brown (2011), who argues that the effects to a large extent are caused by a different socioeconomic composition of those conceived during the pandemic, due to a social gradient in the drafting of army recruits. Here, we test the fetal origins hypothesis, controlling for the individual's socioeconomic origin, using longitudinal micro-level data from the Scanian Economic Demographic Database (SEDD), containing rich economic and demographic information on a population in southern Sweden from the beginning of the nineteenth century up until 2011. We follow birth cohorts born between 1912 and 1923 until 1968. We also test the *infancy inflammation hypothesis*, that scarring due to infections in the first years of life cause health problems in later life (Bengtsson and Lindström 203; Lieuba 2003; Finch and Crimmins 2004). We find support for the fetal origins hypothesis, as prenatal exposure to the 1918 pandemic significantly affected income attainment in adulthood. The effect is, however, rather small compared to the effect of inflammations in the first year of life.

Paper presented in Session *249":* Pathways to health: direct and indirect effects of early life conditions on later health, IUSSP Conference, Busan, 25-31 August, 2013

Introduction – how prenatal and early childhood conditions influence later-life health and socioeconomic performance

Epidemiologists and demographers studying national patterns of the great mortality decline in Great Britain, as well as in Sweden, in the 1920s and 30s found that birth cohorts with improved health in first years of life enjoyed health benefits throughout their lives (Derrick 1927; Kermack, McKendrick, and McKinley 1934). Similar results were later found for the three largest urban départements of France (Preston and Van de Walle 1978), in an extended analysis of Sweden (Fridlizius 1989), and for England, France, Sweden, and Switzerland (Finch and Crimmins 2006). The emphasis on cohort factors in explaining the mortality decline has, however, been questioned based on analyses of even larger data sets. Indeed, in analyzing 28 countries, period factors have been shown to have dominated the decline (Kannisto 1994; see also Barbi and Vaupel 2005). Furthermore, mortality analyses at a county or country level do not provide answers to questions regarding causality, but rather serve as a starting point. For example, based on findings for urban France, the hypothesis that improvements in early-life exposure to disease resulted in improved physical growth and development and to mortality reductions at later ages was put forward (Preston and van de Walle 1978:275).

Studies on Norway further advanced this type of analyses by linking county level infant mortality rate to the risk of dying from arteriosclerotic heart disease at older ages (Forsdahl 1977). The same approach resulted in similar conclusions for England and Wales (Barker and Osmond 1986a). For a later time period, a strong geographical correlation was also found between infant mortality from bronchitis and pneumonia, and death rates from chronic bronchitis and emphysema (Barker and Osmond 1986b). In addition, geographical differences in maternal mortality in England and Wales correlate closely with death rates from stroke in later life, thereby bringing prenatal factors into the analyses (Barker and Osmond 1987). Although these studies are suggesting causal relationships, there is no control for confounders. Another weakness is the lack of the identification of critical periods in the individual's physiological development. Using averages of conditions during a decade or more, it becomes impossible to distinguish between the importance of prenatal, early natal and childhood factors for health in later life.

From the 1940s onwards, medical evidence suggested that the child was more vulnerable during the fetal stage than was previously thought. It was shown that children of mothers diagnosed with rubella during the first part of the pregnancy faced high risks of severe health problems (Gregg 1941; see Dunn 2007). It was also found that the drug Thalidomide, prescribed to women with morning sickness, could cause severe birth defects. Similarly, mothers with alcohol or drug problems disproportionately gave birth to children with congenital defects (see Almond and Currie 2011). As a result, the question was posed whether other, seemingly less severe, conditions during the fetal stage could cause problems that perhaps not showed up directly but, instead, later in life. The literature testing the *fetal origins hypothesis* (Barker 1992) has since grown considerably.

One indicator of fetal stress is birth weight and while this indicator has been criticized, it is still commonly used. Another option is to analyze the health of pregnant mothers (as in Barker and Osmond 1987). One problem relating to both approaches is to find detailed data covering very long time periods, in order to allow for latent diseases of the child to manifest themselves. The often used birth cohort studies from England and the US, for example, only dates back to the mid-twentieth century and can therefore only be used to analyze premature adult mortality, not mortality at older ages (Kuh et al 2009).

Studies using longitudinal individual-level data for 18th- and 19th-century southern Sweden accounted for both prenatal and early post-natal factors (Bengtsson and Lindström 2000, 2003). Macro-level factors experienced by a cohort during early-life were used as a sort of quasi-experiment. Using aggregated measures of maternal mortality during the prenatal stage and infant mortality during the year of birth as indicators of the disease burden, and food prices during pregnancy and in the year of birth as an indicator of the nutritional status of the mother/child, they found that the disease burden in the year of birth played a dominant role in predicting mortality at older ages. Furthermore, they showed that the results were driven by exposure to infectious disease, particularly small-pox and whooping cough, and its influence on respiratory and chest diseases at old ages. In these analyses, key confounding factors, such as socioeconomic status were taken into account. A suggested possible mechanism was that inflammations, caused by the infections, lead to arterial sclerosis which subsequently increases the mortality risk through a number of diseases, including respiratory disease but also coronary heart disease and stroke (Bengtsson and Lindström 2003; Liuba 2003).

In a follow-up study with data up to 1968, whooping cough was found to exercise a strong negative influence on health throughout life, with selection dominating at younger ages and scarring in adulthood and beyond (Quaranta 2013). Similarly, the consequences of early-life exposure to economic upswings and downturns were analyzed for the Netherlands, showing an effect on health later in life (Van den Berg et al 2006, 2009). The advantage of analyzing effects of time-varying exogenous factors in year at birth on individuals is obvious, since effects of repeated events, like the exposure to disease in early childhood, can be tested. A disadvantage is that the time of exposure was unable to fully separate between prenatal and early post-natal exposure.

Several studies have found that conditions in early-life have a net impact on various socioeconomic outcomes (see Bengtsson and Mineau 2009; Currie 2009; Almond and Currie 2011). Based on the British National Child Development Survey it was shown that children with low birth weight were less successful in school, less likely be employed at age 33 years and earned less (Currie and Hyson 1999). A later study, examining the same cohort at 42 years of age show similar results (Case et al 2005). Yet another study using the same data source later on finds effects of birth weight on socioeconomic status at 42 years (Palloni et al 2009). Studies for California (Currie and Moretti 2005) and Norway (Black et al 2011) again show that low birth weight have significant effects on education and labor market outcomes.

While the effects are significant, still they are often very small, similar to effects of birth weight on mortality in heart diseases (see Christensen 2006). Studies of California, show that low birth weight is associated with a few months less education (Currie and Moretti 2007; Johnson and Schoeni 2007; Royer 2009). Effects of low birth weight on earnings seem to be larger (Johnson and Schoeni 2007). Even so, these findings have been questioned, since the design of these studies often does not allow for causal interpretations.

Studies focusing on effects of external shocks due to famines, wars, or other crises attempt to avoid this criticism. Such studies either focus on repeated events, such as the effects of annual variation of food prices (Bengtsson and Lindström 2000, 2003), GDP/capita (van den Berg et al 2006), post-natal disease load using the (local) infant mortality rate as the indicator (Bengtsson and Lindström 2000, 2003; Bengtsson and Broström 2009; Helgertz 2011) *or* on specific events. An example of the latter is a study of the Spanish influenza to test the *fetal origins hypothesis* on adult economic outcomes for the U.S. (Almond 2006). The main

finding was that prenatal exposure to the 1918 influenza pandemic had large negative effects on adult incomes (Almond 2006:708). These findings have, however, recently been questioned, since this effect to a large extent was caused by a change in the socioeconomic composition of new-born due to a social gradient in the drafting of army recruits (Brown 2011, see also Almond and Currie 2011).

In this study, we apply a similar approach to that of Almond, focusing on the effects of the Spanish influenza on socioeconomic performance in adult ages. More specifically, we analyze its influence on the income attainment of males during adulthood. The paper differs, however, in two, arguably important, respects from the work of Almond. First, we use longitudinal individual-level data from the time before the Spanish influenza and onwards, until 1968 to test the *fetal origins hypothesis*. This data, which comes from a rural/semi-urban area in southern Sweden, includes information on individuals' socioeconomic origin as well as on their own labor market career, here measured in terms of income attainment. Secondly, we also test the *infancy inflammation hypothesis*.

The spread of the Spanish flu

Globally, the Spanish influenza is likely to have claimed between 50 and 100 million lives during its short presence in 1918-20 (Johnson & Mueller 2002). In Sweden, the official number of deaths due to the disease amounted to 34,000, killing more than two per cent of those reported infected, compared to 0.1 percent during the typical seasonal flu (Åman 1990). It peaked late in 1918, when the number of deaths was 35 per cent higher than the year before. Apart from being very aggressive, the Spanish influenza has an atypical age-pattern, which was w-shaped. This means that among those who died, a disproportionately large share were young adults, possibly due to an overreaction in the immune system, and possibly due to a partial immunity among the elderly, resulting from the Russian flu of 1889-90. Another characteristic of the influenza pandemic was that pregnant women were particularly vulnerable (Gibson 1919). Apart from its immediate lethal effects, permanent damage from exposure to the Spanish influenza was observed quite early on, for example regarding mental health.

On Midsummer's eve 1918, a worker infected with the Spanish influenza arrived from Kristiania (today's Oslo) to visit his relatives in Hyllinge, a small coal-mining village in southern Sweden, situated only 20 km away from our study area. In a few days, more than 50 people showed symptoms of the influenza. A few weeks later, the influenza had reached all parts of Sweden, transmitted from a number of countries including Norway, Denmark, Germany, Austria, and England (Åman 1990:42). One of the first deaths occurred at the hospital of Malmö, a port town in the very south of Sweden, on June 29, 1918. While the death toll in this first wave was low, the second and very violent wave of the Spanish influenza peaked in October and November 1918, with about 160 deaths per 100 000 inhabitants (Figure 1). A third, comparatively mild, wave of the flu affected Sweden during the spring of 1919.

- Figure 1 here -

Only nine percent of the Swedish population was officially reported as having been ill from the influenza pandemic (Åman 1990:58). This figure is, however, far too low according to contemporary authorities, supported by local and workplace surveys resulting in much higher figures. In some areas, about 75 per cent of the population showed clear symptoms of the influenza, and records from workplaces in southern Sweden and Stockholm suggest that up to 75-80 per cent of all workers below the age of 40 years were afflicted (Åman 1990:58-59).

At Höganäsverket, a coal mine and factory, about 50 kilometers from our study area, a survey conducted by the head occupation health physician in 1918 showed that 61.3 percent of all males and 50.3 percent of all females, families included, was reported as having clear symptoms of the Spanish influenza (Alling 1919: Bilaga I). Among young men in the ages 15-20, the sickness rate was 80.7 percent. No noteworthy differences emerge between white and blue collar workers, similar to figures reported elsewhere in Sweden (see for example Gibson 1919).

Among the mine workers at Höganäsverket, 208 out of the 371 men, 56 percent, were on sick leave for a total of 2,939 days, translating to an average of 14 days per sick worker (Alling 1919: Bilaga II; Höganäsverkets arkiv 1919). Among those reported ill, one third suffered badly from the disease. Assuming that those having weak symptoms stayed at home for a week, similar to individuals afflicted by the seasonal flu today, the 71 workers with severe

symptoms stayed at home for 30 days on average, despite the care provided by the occupation health physicians.

The low influenza mortality from the second half of 1919 and in the following years, is a clear indication of the extent to which the second wave permeated the population, resulting in immunity. Furthermore, blood tests in connection to the 2009 influenza (the so-called Swine flu) which, just like the Spanish influenza, was caused by a H1N1-virus show that those born before the Spanish influenza of 1918-1920 were often immune. In fact, the lowest share of those who were diagnosed with the influenza of 2009 was in the age group above 65 years of age (Smittskyddsinstitutet 2010).

The rapid spread of the influenza meant that already by the end of November 1918, two thirds of the total deaths in Sweden due to the Spanish influenza had occurred. By the end of the year, the figure was four fifths (Figure 2). Thereafter, the number of deaths fell, only to re-surge somewhat during the third wave in March-April 1919. Despite the substantial underreporting of persons having influenza symptoms, the morbidity rate showed almost an identical time pattern indicating that neither the diagnosis, nor the treatment improved over time. By the end of 1918, some 83 per cent of those ever-reported ill had experienced the disease. Thus, the time pattern of the influenza appears to be quite clear. The first few cases, of the milder form, occurred around Midsummer's Eve in 1918 in southern Sweden. It then appeared all over Sweden. The second wave, which was much more virulent, started in September and peaked in October-November the same year. As a result, by the end of 1918, the two thirds of the population that would become ill, possibly more, had experienced the disease. In the third wave, during the spring of 1919, the remaining ten percent of the deaths occurred.

- Figure 2 here -

Based on death certificates from the area of study in this article, the first death due to the Spanish influenza occurred on September 19, 1918. The afflicted were either diagnosed with Spanish flu, pneumonia, or both. No one in the sample is reported to have died during the first and milder wave. The spread over time was similar to the country as a whole but with a slightly slower progression: only two thirds out of all who died in the Spanish influenza

before the summer of 1919 had died by the end of 1918 (Figure 2), compared to four fifths for Sweden as a whole.

The use of the Spanish influenza as a quasi-experiment is based on a number of assumptions. First, the event should not be planned for or anticipated. Second, when the event occurs place, there should be no means to avoid it, as well as no treatment. Third, there should not be any other factor that takes place at the same time that could have an influence on the outcome of the event (no omitted variable). While we believe that the first two do not pose any problems, the latter assumption does warrant further examination. Brown (2011) showed that an omitted variable – socioeconomic status at birth - caused Almond to overestimate the consequences of in-utero exposure to the Spanish flu. More specifically, the fact that the recruitment of American soldiers during the First World War was socioeconomically biased, as the authorities avoided recruiting married men of poor families, this changed the social composition of the 1918 birth cohort. However, an in-utero effect from exposure to the Spanish influenza has also been found by Brown (Brown 2011; Almond and Currie 2011). While Sweden was not at war in 1918, more men than normal were nonetheless recruited to the army. The magnitude was, however, different from the situation in the US, and potential recruits could not be excluded based on socioeconomic background. Furthermore, a change in the composition of new-born is primarily a problem if the information on socioeconomic status at birth is omitted. This is not a problem in the present study, as we have access to information on parental occupation.

Data and methods

The data used in the analysis comes from the Scanian Economic Demographic Database (SEDD, version 3.1, Bengtsson et al 2012). The database contains longitudinal information on all individuals residing in five parishes in southern Sweden. Individuals are followed over time using data obtained from continuous population registers, containing information on demographic events, including migration to and from households. The information in the registers has been checked against the birth and death registers to adjust for possible under-recording of events in the population registers. The data from the population registers have further been linked to poll-tax registers (*mantalslängder*), providing annual information on

the occupation of the family head up to 1945. In addition, individual-level occupational information from annual income- and taxation registers has been added from the 1860s and onwards, and individual incomes from 1903 onwards. The resulting database contains all individuals living in the five parishes, either from birth or from in-migration until they die, out-migrate or the study period ends.

The sample selected for this paper consists of all males born between 1912 and 1923, ever observed in five parishes, approximately 1,700 individuals. These are allocated into birth cohorts, depending on at what age they were exposed to the influenza pandemic. Based on data national on morbidity and mortality, the influenza period is defined as from September 1918 to and including April 1919 (see shaded area in Figure 2).

The mechanisms that are investigated in this paper direct a particular focus towards two distinctive phases during the individual's development. As previously outlined, the *infancy inflammation hypothesis* postulates that individuals who are exposed to infectious disease or other environmental insults during their first year of life could experience a chronic inflammatory response, thereby elevating the risk of experiencing a range of diseases throughout life and resulting in a varying degree of debilitation. In analyzing the outcomes of individuals subjected to this form of exposure, individuals who were born before, but no more than a year before, September 1918 were selected. Given the expectation that the influenza virus during the period September 1918 and until April 1919 virtually permeated the entire population, those within this group should at some point in time during their first year of life have been exposed to the influenza virus.

To test the *fetal origins hypothesis*, individuals who were exposed to the influenza during the fetal stage were selected, consisting of those born after, but within nine months of end of the influenza period. Another group that was affected by the influenza during *both* of the aforementioned sensitive periods was identified by selecting those born *during* the influenza epidemic. More specifically, individuals belonging to this group consist of those who were exposed both in-utero and during infancy.

All birth cohorts are presented in Table 1, in which the reference category used in the multivariate analysis is represented by individuals born over nine months after the end of the influenza period and until 1923. Thereby, this group represents individuals who were

conceived after the pandemic had subsided and should therefore be unexposed to the influenza.

- Table 1 here -

Non-random selection of individuals into the study sample must be carefully examined. More specifically, those having been exposed to the influenza pandemic during any of the sensitive periods may be characterized by certain important unobserved characteristics. For example; if exposure to the influenza caused the culling of the weakest fetuses, those surviving in-utero exposure to the pandemic could be positively selected, potentially resulting in a misinterpretation of the true consequences of influenza exposure. Such selection during the fetal stage could show up as a change in the sex-ratio at birth, since males are more vulnerable to in-utero disturbances than females. Also, a high frequency of still-births indicates an unfavorable development in utero.

Sex ratio at birth among all recorded births in the SEDD and for the cohorts of interest for this article was normal, with 104 male for each 100 female live births, as shown in Table 1. The same conclusion applies to the share of (singleton) stillbirths. Among cohorts born during the influenza pandemic and therefore exposed during the last trimester of pregnancy, the sex ratio is lower, around 47 percent male births. Combined with the elevated share of stillbirths, the lower sex ratio would appear to be linked to an increased incidence of in-utero mortality, particularly among male fetuses. While one should be cautious in drawing too far reaching conclusions due to small numbers, the figure could be indicative male fetuses being more vulnerable than female fetuses, and therefore failing to survive past the fetal stage to a greater extent. The same tendency, albeit less pronounced, can be observed among those exposed during the earlier phases of pregnancy and born within nine months of the end of the pandemic (exposed both in-utero and during infancy), as indicated by the 49 percent male births observed. The observed sex-ratios for these birth cohorts could hence be indicative of a certain degree of positive selection, though very modest, into the sample among those exposed while in-utero. Furthermore, the share of still births does not suggest any elevation in the risk of not surviving the fetal stage.

Table 1 also suggests that the birth cohorts were characterized by quite substantially improving survival chances past infancy and childhood occurring over time. Consistently, the

most elevated infant or child mortality is experienced by those who would have turned five before the time of the pandemic outbreak (born 1912-1913). Almost consistently, subsequent birth cohorts are observed with higher shares of survivors, with cohorts exposed to the pandemic during sensitive periods frequently being observed with 10 percentage points higher survival shares than those turning five prior to the influenza outbreak, not being exposed to the influenza during a sensitive period. The observed pattern is reinforced by the parishes' infant mortality rate, displayed in Figure 3. The two influenza years, 1918 and 1919, are highlighted, clearly characterized by distinctly lower infant mortality rates than the preceding years, despite the pandemic. Consequently, neither Table 1 nor Figure 3 suggest that individuals exposed during infancy or in-utero suffered from an elevated mortality risk in early childhood that could have been indicative of the existence of positive selection.

- Figure 3 here -

The outcome variable is defined as the natural logarithm of the individual's annual taxed income, adjusted by means of the consumer price index. The income information was manually gathered from the records of the local tax authorities, typically at the parish level. Starting from the income year 1902, Sweden was among the first nations to implement a comprehensive taxation system that relied on the individual's self-reporting of various types of incomes, including income from capital, labor and self-employment. Until 1943, the obligation to report one's incomes was, however, contingent on the accumulated income exceeding a certain threshold, hovering between 500 and 600 SEK. For example, incomes earned during 1902 the threshold amounted to 500 SEK, which in 2012 would be the equivalent of approximately 27,000 SEK. After 1943, all positive incomes were subject to reporting to the tax authorities which at this time had become considerably more centralized. A potential problem is, however, presented by the eventuality that individuals who were exposed to the influenza pandemic during sensitive periods to a greater extent only are characterized by a marginal labor market attachment, manifested by a failure to achieve an income exceeding the aforementioned threshold. Looking at all males observed in adulthood, Table 1, however, reveals that only a very small share was never observed with a positive income, and with negligible differences across birth cohorts. More specifically, between 96 (exposed during infancy) and 99 (between 1-2 years of age at exposure) percent were at some point between the ages 20 and 45 observed with a positive income.

The living standards improved considerably during the time period examined in this paper, particularly after the Second World War. Figure 4 shows that the mean real annual income among all working age males (20-60) in the study area tripled over the time period. Although this increase may overestimate the magnitude of the *actual* living standards improvements, it clearly illustrates the rapid societal transformation occurring at the time. Real wages for industrial workers in Sweden as a whole shows the same development.

- Figure 4 here -

The rapidly improving labor market from the 1940s is further confirmed by Figure 5, displaying the country-wide unemployment rate for union workers. The earlier birth cohorts, entering the labor market during the early 1930s, suffered through the aftermath of the depression, though less severe in Sweden as in the US, with unemployment rates steadily exceeding ten percent. Coinciding with the rising real wages and real incomes, the unemployment level rapidly dropped to reach around two percent at the end of the 1940s, where it stayed throughout the period examined. Thus, while the sample analyzed in this article only covers 12 consecutive annual birth cohorts, it would appear plausible to expect the rapidly changing macroeconomic environment to exercise an important influence on the outcome of interest for this article.

- Figure 5 here -

Thus, the economic boom occurring from the early 1940s and throughout the follow-up period of this paper could have important consequences for the study population. Assuming a typical age at labor market entry around 20 years of age, the earliest cohort examined in this article spent their first 5-10 working age years in a considerably less favorable labor market environment than the cohort born after 1920, entering the labor market during the boom of the 1940s.

In more carefully exploring this possibility, Figure 6 indeed suggests that the favorable conditions during the 1940s and beyond particularly favored the later born cohorts, including those exposed to the pandemic during infancy/in-utero. The pattern is even more evident in Figure 7, showing cohort wages. Individuals conceived after the pandemic enjoy a quite steep earnings growth during their 20s, clearly not visible among those born during the early 1910s.

Inconsistent with the hypotheses of the paper, neither of the cohorts having been exposed to the influenza pandemic during a sensitive period hence emerge as particularly disadvantaged.

- Figures 6 and 7 here

The dataset extends through the year 1968, and individuals are followed on a yearly basis, conditional on having a positive income. We exploit the panel character of the data, estimating models by means of OLS regression, controlling for unobserved heterogeneity through individual random effects. Variable means are presented in Table 2. The unemployment rate is included in the models in order to capture a key labor market demand factor, as suggested by the descriptive analysis. The infant mortality rate (IMR) during the year after the birth of the individual is also included to capture the possibility that the environment facing individuals shortly after birth – independent of the influenza pandemic – affects their future income attainment. This variable is operationalized both as the unadjusted IMR, and as a dichotomous variable, indicating whether the IMR during the year subsequent to the individual's year of birth exceeded the long term trend by more than ten percent.

- Table 2 here -

Results

The article tests the fetal origins hypothesis by estimating the effects of exposure to the influenza pandemic of 1918/1919 on income attainment in adulthood. Model 1, presented in Table 3, includes only birth cohorts and controls for age and age squared, providing a first indication regarding how the income attainment of birth cohorts differ depending on their exposure to the influenza, controlling for the age at which the income was earned. Consistent with the rapidly increasing real incomes in Sweden at the time, those born later almost consistently enjoyed higher incomes. Furthermore, the magnitude is far from negligible. Holding age constant, those born in 1913 or before and therefore exposed to the pandemic after the age of 5 earned an annual income amounting to no more than 68 percent ($e^{-0.382}$) of those conceived after the influenza period.

- Table 3 here -

The generally improving incomes of individuals belonging to later birth cohorts is consistent with the increasingly beneficial labor market conditions facing later born individuals, possibly also combined with the improving health environment facing individuals during their first year of life and beyond. The cohort exposed to the influenza pandemic while in-utero, however, represents the possible exception to an otherwise quite consistent pattern. More explicitly, and similar to the results of Almond (2006), this group is observed with a slightly accentuated income disadvantage, amounting to about one percentage point, compared to the group born immediately prior, exposed both while in-utero and during infancy. Compared to the subsequent birth cohort, conceived after the end of the influenza pandemic, those exposed in-utero are predicted to attain an income being about 11 percent points lower.

In Model 2, the individual's origin social class as well as whether they were born in their parish of residence is added. As expected, the transmission of socioeconomic resources across generations strongly exerts an important influence on an individual's labor market outcome. Compared to the reference category, represented by individuals whose father worked in a skilled occupation, individuals with unskilled fathers are predicted to experience approximately ten percent lower incomes. A similar disadvantage is observed among individuals for whom the social class of their fathers is unknown, suggesting that this group de facto most likely being dominated by individuals from lower class origins.

The birth cohort estimates are only affected marginally by the inclusion of discussed parameters. More specifically, the earliest birth cohorts are still observed with considerably lower incomes than the later birth cohorts, holding everything else equal. Again, a slightly accentuated disadvantage is experienced by those only exposed to the influenza pandemic while in-utero, similar to the previous specification in terms of magnitude. The size of the disadvantage should not, however, be exaggerated, as the parameter fails to be significantly different from the immediately preceding birth cohorts, exposed both while in-utero and during infancy and only during infancy, as suggested by the 90% confidence intervals.

In Models 3 and 4, information is added to control for the importance of changes over time in the health environment facing individuals shortly after birth, consistent with the *infancy inflammation hypothesis*. More specifically, the models control for the parishes' infant

mortality rate (IMR) during the year after birth, in order to isolate the time of exposure to occur after birth and not during the fetal stage. In Model 3, IMR enters the model in an unadjusted manner, suggesting that for each one unit increase in the IMR during the year after birth, the individual's attained adulthood income decreases by 0.1 percent. Comparing the earlier birth cohorts included in the sample, exposed to environments with an IMR of around 100 deaths per 1000 live births, the last birth cohorts, subjected to an IMR of around 50‰, the latter cohort is hence predicted to enjoy an income advantage amounting to about 5 percent. Consequently, during the period preceding the influenza pandemic, when year-to-year variation in IMR amounting to 50 units were not uncommon, the circumstances experienced during the first year of life could, according to the results, have non-negligible consequences for the individual's labor market outcome.

Including IMR in the models slightly moderates the disadvantage experienced by the earliest birth cohorts, exposed to considerably more adverse health conditions during the year after birth. Despite this, the earliest birth cohorts still experience a substantially lower income attainment according to Model 3. Turning to the cohorts exposed to the influenza pandemic during sensitive periods, the slightly increased disadvantage for those exposed while in-utero, disrupting the otherwise consistent pattern, still remains.

Due to the secularly declining trend over time in the IMR, yet less evident here due to the limited number of consecutive annual birth cohorts included, it may be difficult to identify a meaningful effect. Therefore, Model 4 separates the short-term cyclical variation from the long-term trend in the IMR, identifying years characterized a cyclical variation exceeding the trend by at least ten percent. This is interpreted as years characterized by an elevated disease load, or otherwise worsened health conditions facing individuals during their first year of life. The estimates from Model 4 suggests that individuals exposed to an IMR exceeding the long-term trend by at least ten percent experience an income attainment in adulthood that was about five percent lower than someone not exposed to this insult. Remaining parameters again remain largely unaffected.

As suggested by the preceding analysis, the existing macroeconomic conditions could exercise an important influence on the individual's attained income. Models 5 and 6 are therefore extended to control for the annual unemployment rate, as well as the area in which the individual resides. More specifically, the parishes Hög and Kävlinge industrialized earlier

and to a greater extent than the parishes Halmstad, Kågeröd and Sireköpinge. The greater career opportunities within the expanding industry is arguably also reflected by the estimates, where those residing in Hög and Kävlinge enjoy around 23 percent higher incomes, all else equal. The effect obtained for the unemployment rate is also by no means negligible, suggesting that every one unit increase is associated with a three percent decrease in the individual's income.

Whereas previous model specifications have only marginally affected the birth cohort estimates, controlling for labor market demand factors indeed suggests that a substantial part of the disadvantage experienced by the earliest cohorts can be attributed to these groups being subjected to considerably worse labor market conditions. More specifically, the baseline earnings of the first birth cohort amounting to 68 percent of the reference category's according to Model 5 amounts to 83 percent. Controlling for labor market factors does not, however, alter the pattern previously observed for the birth cohorts. Hence, those exposed to the influenza while in-utero appear to experience a slightly accentuated disadvantage. The estimated parameter for individuals exposed both in-utero and during infancy is now, however, not statistically significantly different from the reference category, unexposed to the influenza pandemic. Recalling the indications obtained from the data and methods section, this could be evidence of this group being subjected to a certain degree of pre-natal culling (through the low ratio of male births), and the individuals in this group thereby being positively selected. While hardly being conclusive evidence, this could explain the improved outcome observed among this group. The effect from the IMR level remains unaltered compared to Model 3, as well as statistically insignificant.

Lastly, Model 6 uses the threshold definition of the IMR during the year after birth, as opposed to the continuous specification used in Models 3 and 5. The results again suggests that exposure to adverse conditions during the year after birth is associated with a lower income attainment, slightly accentuated compared to Model 4, estimated without labor market characteristics. Here, exposure to health conditions characterized by an IMR exceeding the trend by at least ten percent is associated with an almost six percent lower income. Figure 8 displays the birth cohort effects obtained from the gradual extension of the specifications discussed. Arguably, the overestimation of the earnings disadvantage experienced by the earlier birth cohorts that the results from failing to control for labor market demand factors emerges even more clearly from the figure. As regards the cohorts exposed to the influenza

pandemic during what has defined as sensitive periods, the parameter estimates again indicate that those exposed while in-utero appear to experience a slight disadvantage. Again, the statistical insignificance of the estimate for individuals exposed both while in-utero and during infancy could suggest that the group in fact does not differ from the reference category.

- Figure 8 here -

Conclusions

To this date, the study of the socioeconomic consequences of exposure to the Spanish influenza remains few within the social sciences. Despite the approaching centennial of its outbreak, its relevance as a topic for academic study is, however, undisputed. The interest in the pandemic due to its devastating global health consequences is not, however, limited to historical studies. Instead, the last decade has been characterized by a growing public awareness regarding the potential consequences of a new global influenza pandemic. Not only did the fears of another pandemic that dominated the news headlines during the 2009 Swine flu outbreak introduce the abbreviation H1N1 into the global population's vocabulary. A global influenza outbreak and its consequences were also key in the 2011 Steven Soderbergh film Contagion. While today's society surely would fight a new influenza pandemic with different means than in 1918-1919, understanding its long-term consequences for those affected remains of great relevance. Using longitudinal individual-level data on the entire population of five parishes in southern Sweden, this paper presents a contribution to the improved understanding of the consequences of exposure to the Spanish flu.

While the sample examined in the paper, combined with detailed information regarding the timing of the spread of and exposure to the influenza, offers unique possibilities in analyzing the consequences of influenza exposure, its weaknesses should not be ignored. Firstly, the number of individuals that constitute the birth cohorts of particular interest in addressing the research questions at hand is quite small. Secondly, the amount of migration to and from the parishes was substantial. The results do strongly indicate that being a migrant mattered to an

individual's socioeconomic attainment, but it remains beyond the scope of this paper to fully elaborate on the interaction between migrant status and influenza exposure.

Despite the small sample used for the paper, the results are nonetheless highly interesting and relevant due to data quality. Firstly, in analyzing the consequences of exposure to the Spanish influenza, the paper represents the first attempt at following individuals' income attainment *over time*, as opposed to using a cross-sectional approach. Another important contribution is represented by the important influence on the outcome from the labor market conditions during this time of economic expansion, previously largely unexplored in micro-level studies on historical contexts.

Overall, we find support for the *fetal origins hypothesis*, that exposure to the pandemic during the fetal stage had a negative impact on adulthood income attainment, but the magnitude of the effect is weak. This is in line with the findings of Almond and Brown, whose results from U.S. census data suggest that the group that was exposed in-utero is associated with adverse adulthood outcomes. With effects that are quantitatively more important, we also find support for the *infancy inflammation hypothesis*; that those exposed to an adverse health environment during the year after birth experience lower incomes in adulthood, consistent with previous findings by Bengtsson and Broström, Helgertz and Quaranta.

Acknowledgements

Jonas Helgertz is grateful to the post-doc project *Early-life health conditions and adulthood socioeconomic outcomes in Sweden, 1968-2001*, financed by Swedish Council for Working Life and Social Research (FAS, grant nr 2011:1338). Both authors are appreciative for research resources provided by the Linnaeus Centre for Economic Demography (CED), at Lund University funded by Swedish Research Council (VR, grant nr. 2006:79/2008:6588).

References

Alling, Gustav (1919). Översikt över "Spanska sjukans" förlopp bland Höganäsverkets personal. *Allmänna Svenska Läkartidningen, 25:581-589.* Stockholm.

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. *The Journal of Economic Perspective.* 25, 3:153-172

Åman, M. 1990. *Spanska sjukan. Den svenska epidemin 1918-1920 och dess internationella bakgrund.* Uppsala: Almqvist & Wiksell International Stockholm

Barker, D.J.P. and bd, C. 1986a. "Infant Mortality, Childhood Nutrition, and Iscaemic Heart Disease in England and Wales", Lancet 327 (6998): 171-74.

Barker, D.J.P. and bd, C. 1986b. "Respiratory Infection And Adult Chronic Bronchitis In England And Wales", *British Medical Journal* (Clinical Research Edition), Vol. 293, No. 6557, Nov. 15: 1271-1275.

Barker, D. 1994 *Mothers, babies, and disease in later life.* London: BMJ Publishing Group.

Bengtsson T., Broström G. 2009. "Do Conditions in Early Life Affect Old-Age Mortality Directly and Indirectly? Evidence from 19th Century Rural Sweden." In Bengtsson T, Mineau G.P. (eds.) Early Life Effects on Socio-Economic Performance and Mortality in Later Life. A Full Life Course Approach Using Contemporary and Historical Sources. Special Issue of *Social Science & Medicine,* 68(9), 1583-1590.

Bengtsson T., Lindström M. 2000. Childhood misery and disease in later life: The effects on mortality in old-age of hazards experienced in early life, southern Sweden, 1760-1894. *Population Studies* 54:263-277.

Bengtsson T., Lindström M. 2003, "Airborne infectious diseases during infancy and mortality in later life in southern Sweden, 1766-1894", International Journal of Epidemiology, 32, 286-294.

Bengtsson, T, and Mineau, GP. (eds.) 2009. "Early life conditions and other factors that influence survival and longevity: the use of contemporary and historical sources." Special Issue of *Social Science and Medicine,* 68(9):1561-1658.

Bengtsson, T., Dribe, M., Svensson, P. 2012. Scanian Economic-Demographic Database (SEDD), version 2012:1.

Black, S, Devereux, P, and Salvanes, K. 2007. "From the Craddle to the Job Market? The Effect of Birth Weight on Adult Outcomes of Children." *The Quarterly Journal of Economics,* 122(1):409-439.

Brown, R. 2011. The 1918 U.S. Influenza Pandemic as a Natural Experiment, Revisited. Duke University. Conference paper.

Case, A., Fertig, A, and Paxson, C. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics*, 24(2):365-389.

Currie, J. 2009. Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in childhood, and Human Capital Development. *Jounral of Economic Literature,* 47, 1:87-122.

Derrick, V.P.A. 1927. Observations on (1) errors in age in the population statistics of England and Wales, and (2) the changes in mortality indicated by the national records. Journal of the Institute of Actuaries 58: 117-159.

Finch, CE, and Crimmins, EM. 2004 . "Inflammatory exposure and historical changes in human life-spans." *Science,* 305(5691):1736-1739.

Forsdahl, A. 1977. “Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease?” *British Journal of Preventive and Social Medicine*, 31:91-95.

Fogel, R. 1993. "New sources and new techniques for the study of secular trends in nutritional status, health, mortality, and the process of aging." *Historical Methods* 26(1): 5-43.

Fridlizius, G. 1989. “The deformation of cohorts: nineteenth-century decline in a generational perspective.” *Scandinavian Economic History Review*, 37(3), 3–17.

Gibson, Gunnar 1919. Några iakttagelser över “Spanska sjukan” bland bruksarbetare i Sandviken. *Allmänna Svenska Läkartidningen, 7:160-165.* Stockholm.

Helgertz, J, 2010. “Early-life conditions and income attainment in Sweden during 1968-2010: Using sibling data to further exlpain why country of origin should matter” in Helgertz, J, *Immigrant careers. Why country of origin matters.* Lund: Media-Tryck.

Höganäsverkets archive. 1919. Spanska sjukan vid gruvdriften under 1818, Diverse. RF/P U16. Höganäs.

Johnson, N.P.A.S., and Mueller, J., (2002): Updating the Accounts: Global Mortality of the 1918-1920 “Spanish” Influenza Pandemic. *Bull. Hist. Med.*, 76: 105-115.

Kannisto V., Christensen K., Vaupel J.W. 1997. ”No Increased Mortality in Later Life for Cohorts Born during Famine”, *Am. J. Epidemiol.* 145: 987 – 994

Kermack W., A. McKendrick and P. McKinlay. 1934. ”Death rates in Great Britain and Sweden: some regularities and their significance.” *Lancet*, March, 698-703.

Kuh, D, and Ben-Shlomo, Y. (eds.) 2004. A Life Course Approach to Chronic Disease Epidemiology. Second edition. Life Course Approach to Adult Health Series. Oxford: Oxford University Press.

Kuh, D., Shah, I., Richards, M., Mishra, G., Wadsworth, M., Hardy, R. 2009. "Do economic cognitive ability or smoking behaviour explain the influence of lifetime socio-economic conditions on premature adult mortality in a British post war birth cohort", *Social Science and Medicine,* 69:1565-1573.

Lindhagen, E. 1928. "Grippe und Lungentuberkulose. Mortalitätsstatistische Ergebnisse", *Zeischrift für Tuberkulose,* 46:4, 321-337.

Lindström, M, and Davey Smith, G. 2007. "A Life Course Perspective to the Modern Secular Mortality Decline and Socio-economic Differences in Morbidity and Mortality in Sweden." In Bengtsson T (ed.) *Perspectives on Mortality Forecasting. Vol. V. Cohort factors: How conditions in early life influence mortality later in life*. Stockholm: Swedish Social Insurance Agency, 9-29.

Liuba, P. 2003. *Arterial injury due to infections in early life: A possible link in coronary heart disease*. Lund: Lund University Hospital.

Palloni, A, Milesi, C, White, RG, and Turner, A. 2009. "Early childhood health, reproduction of economic inequalities and the persistence of health and mortality differentials." Special issue of *Social Science and Medicine,* 68(9):1561-1744.

Preston, S., Van de Walle, E. 1978. "Urban French Mortality in the Nineteenth Century", *Population Studies,* 32:2 (Jul.), 275-297

Royer, H. 2009. "Seperated at Girth: US Twin Estimates of the Effects of Birth Weight", *American Economic Journal: Applied Economics 1:1, 49-85.*

SOS Befolkningsrörelsen åren 1918-1920. 1926. Stockholm: Kungliga Statistiska Centralbyrån.

SOS Allmän hälso- och sjukvård år. 1918. Stockholm: Kungliga Statistiska Centralbyrån

SOS Allmän hälso- och sjukvård år. 1918. Stockholm: Kungliga Statistiska Centralbyrån

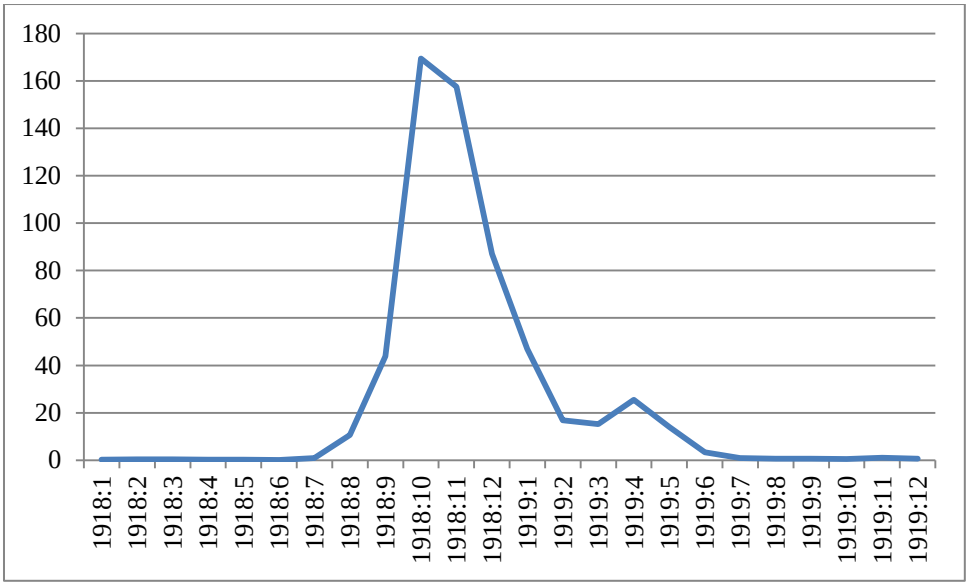
Smith, JP. 2009. “The Impact of Childhood on Adult Labor Market Outcomes.” *Review of Economics and Statistics*, 91(3):478-489.

Smittskyddsinstitutet 2010: Influensarapport, vecka 20 (17/5-23/5), 2010. Retreived from http://smi.se/publikationer/veckorapporter/influensarapporter/sasongen-20092010/influensarapport-vecka-20-175---235--2010/ on 2013-08-13.

Van den Berg, G.J, Lindeboom, M, and Portrait, F. 2006. ”Economic conditions early in life and individual mortality.” *American Economic Review*, 96:290-302.

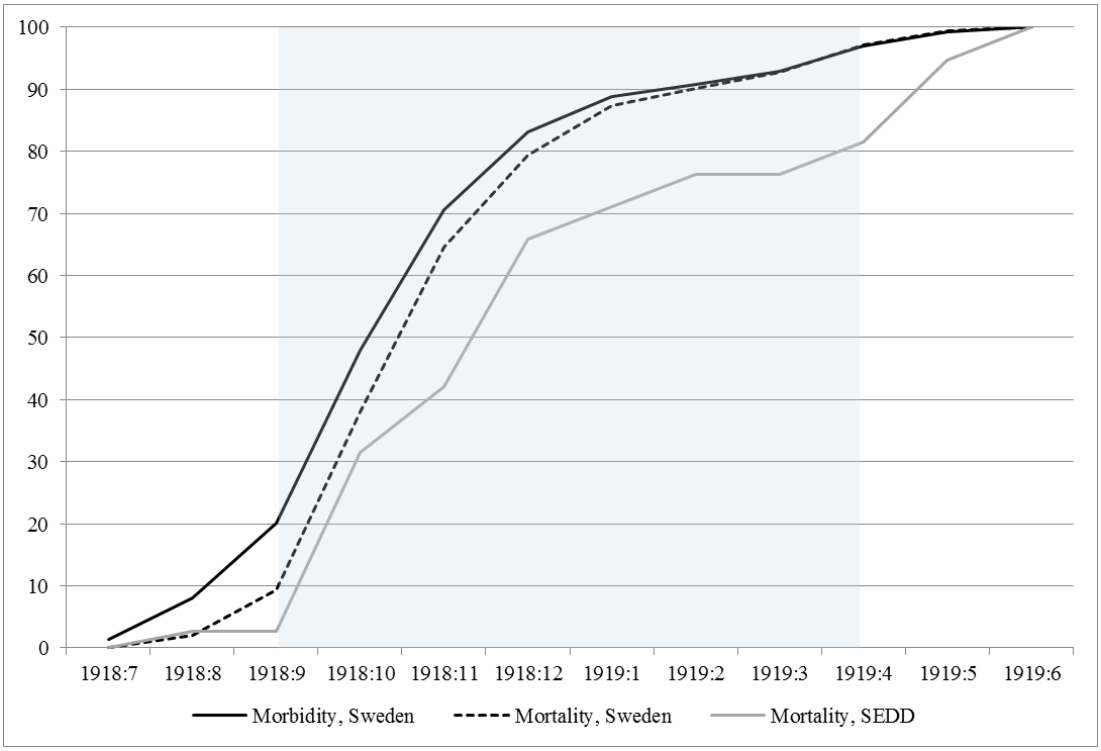
Figures

Figure 1. Influenza mortality in Sweden 1918-19. Per 100 000.



Source: Lindhagen 1926:3

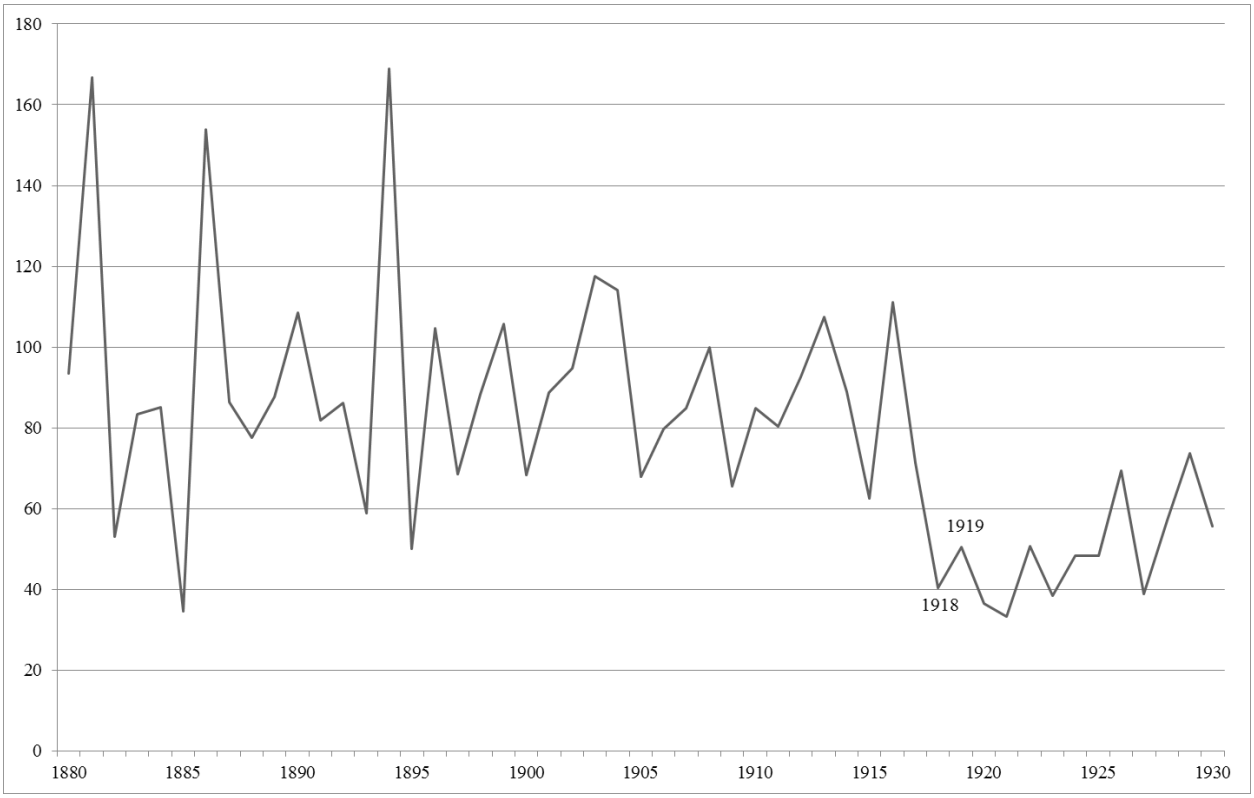
Figure 2. Accumulated percent morbidity and mortality in Spanish flu in Sweden 1918-19.



Sources: Scanian Economic-Demographic Database (SEDD), Bengtsson, Dribe and Svensson 2012. Lindhagen 1926, SOS Allmän hälso- och sjukvård år 1918, 1919 (Table 8), for Sweden

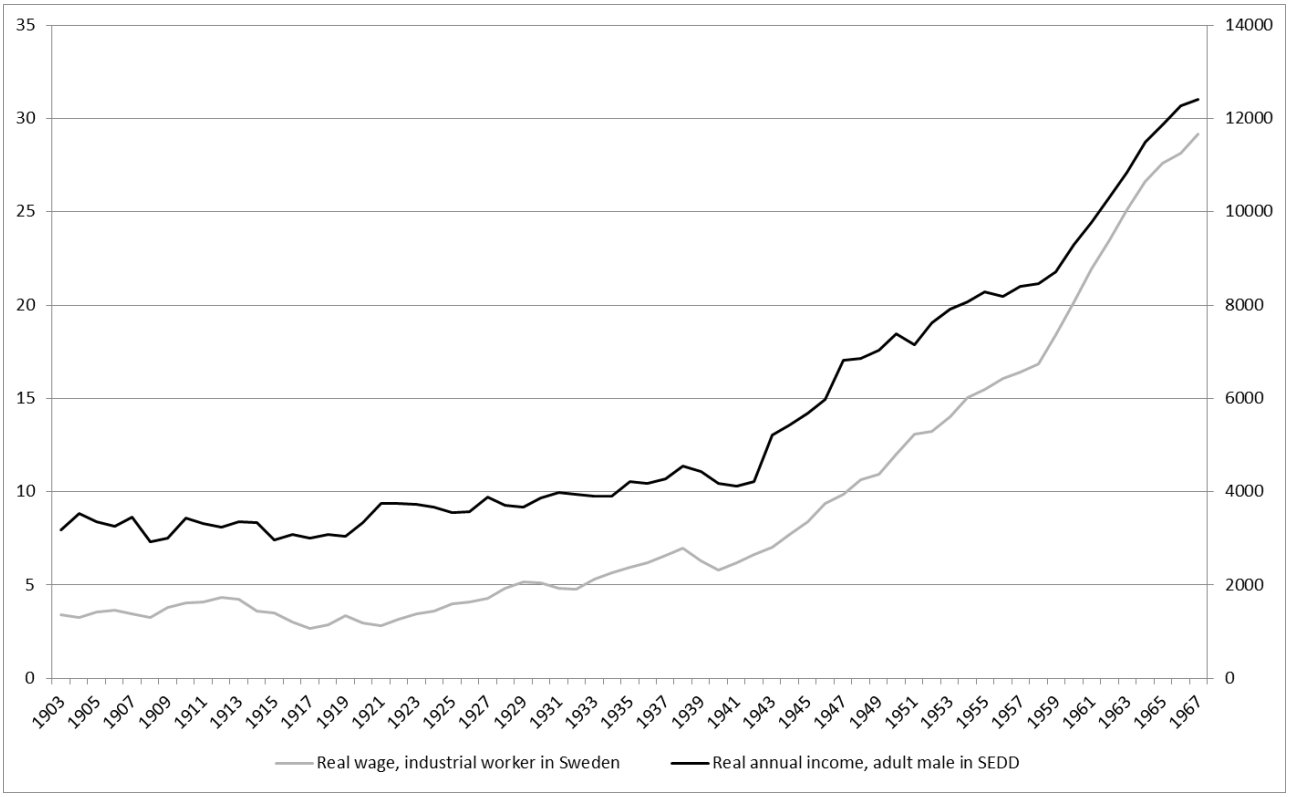
Note: Shaded area represents influenza period

Figure 3: Infant Mortality Rate (deaths per thousand live births), 1880-1930



Source: Scanian Economic-Demographic Database (SEDD), Bengtsson, Dribe and Svensson 2012.

Figure 4: Real wage and real annual income, 1903-1967



Sources: Real wage: Edvinsson: Wages of employees in seven types of activities in Sweden 1850-2000, version 1.0

Annual income: Scanian Economic-Demographic Database (SEDD), Bengtsson, Dribe and Svensson 2012.

Figure 5: Annual unemployment rate, 1911-1968

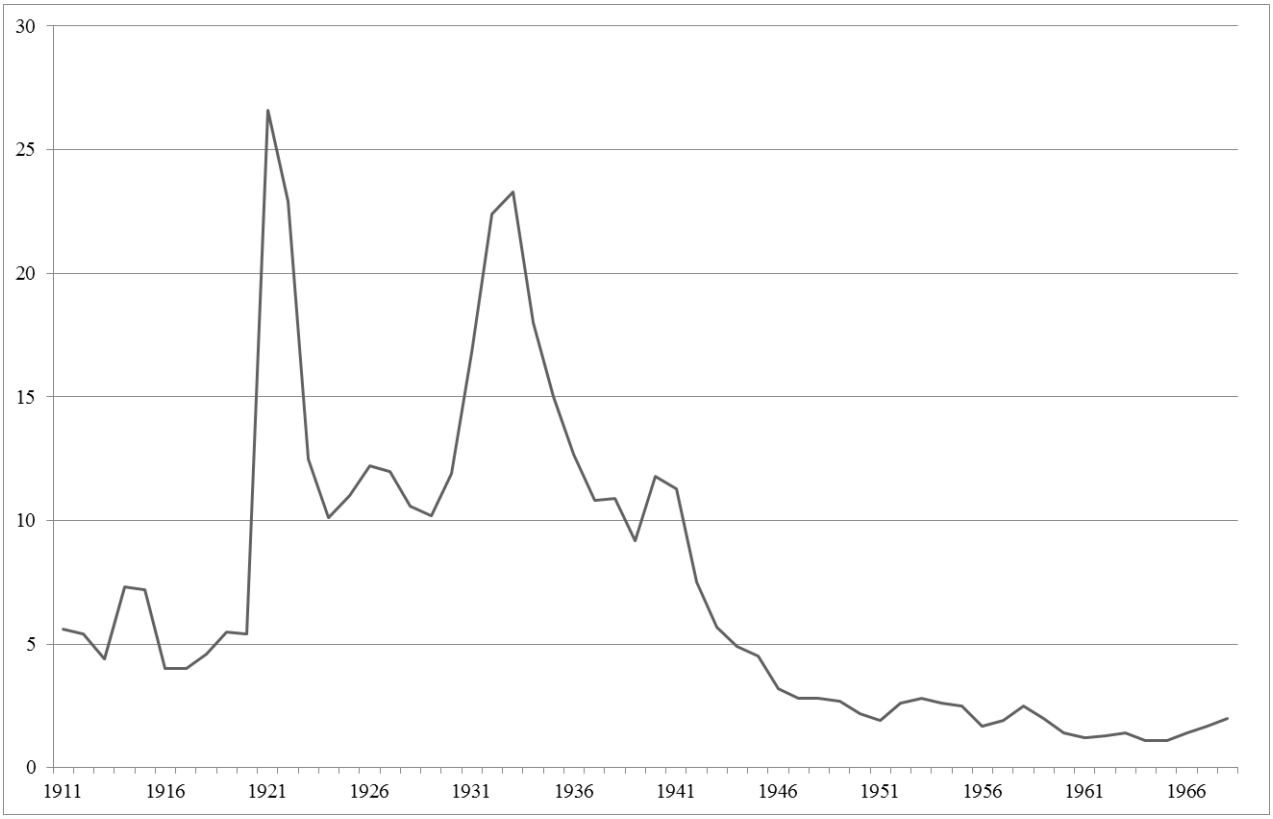
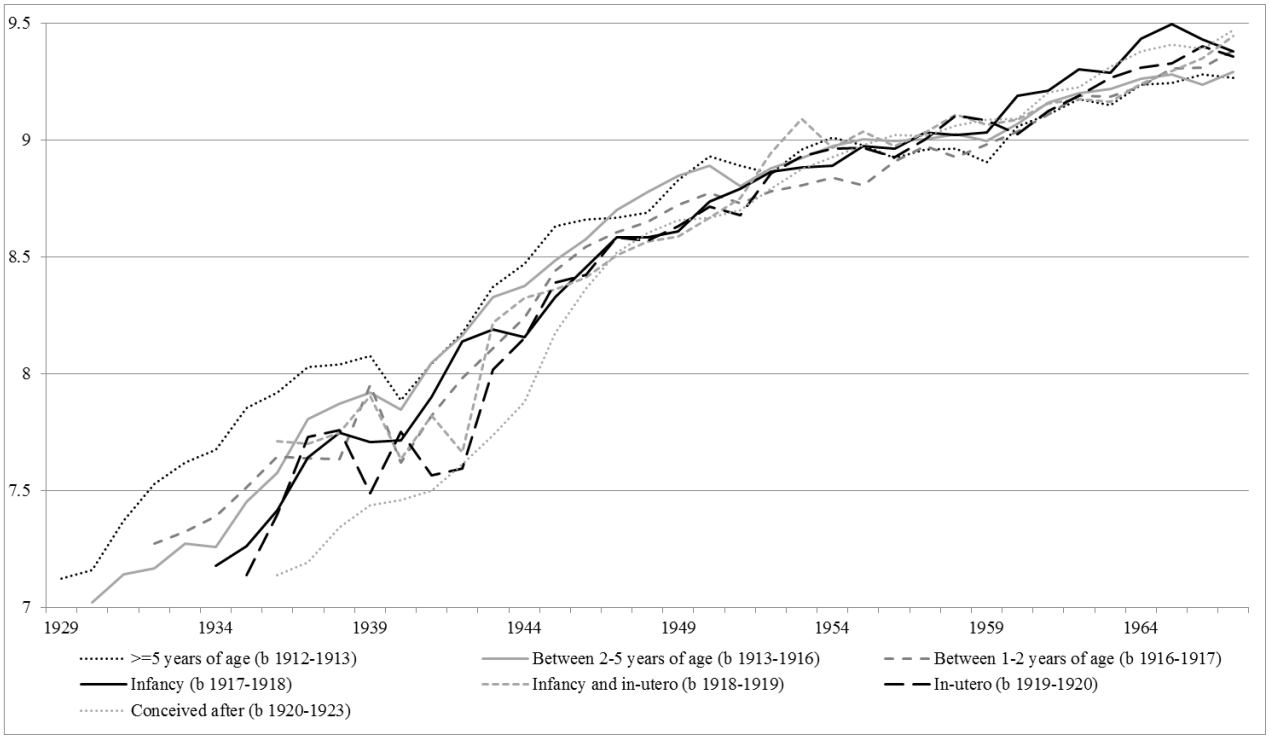
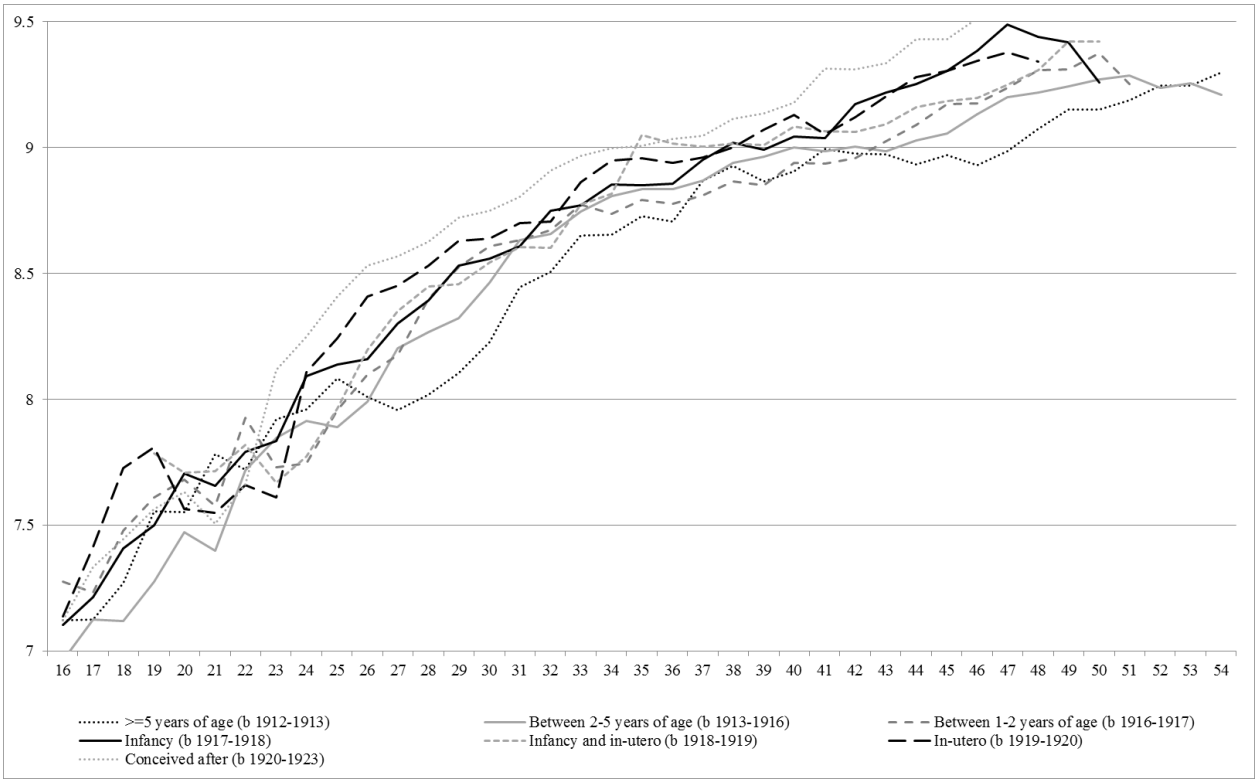


Figure 6: Mean annual real income trajectories, by year and birth cohort



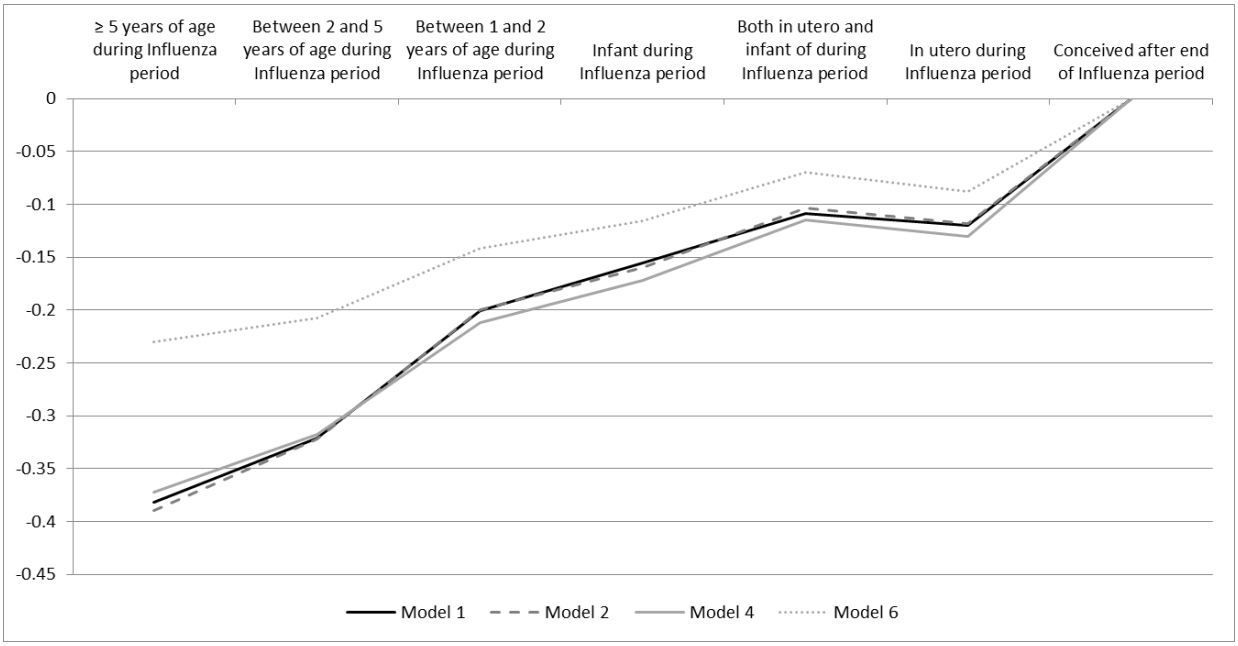
Source: Scanian Economic-Demographic Database (SEDD), Bengtsson, Dribe and Svensson 2012.

Figure 7: Mean annual real income trajectories, by age and birth cohort



Source: Scanian Economic-Demographic Database (SEDD), Bengtsson, Dribe and Svensson 2012.

Figure 8: Birth cohort effects



Tables

Table 1: Birth cohort characteristics

	(1) Individuals	(2) Male ratio	(3) % still births	(4) % surviving past the age of 1	(5) % surviving past the age of 5	(6) % surviving past the age of 15	(7) % observed with positive income between ages 15-45
≥ 5 years of age during Influenza period	244	52.52	1.88	88.37	83.71	82.74	96.85
Between 2 and 5 years of age during Influenza period	402	51.52	1.28	90.57	86.18	83.74	96.55
Between 1 and 2 years of age during Influenza period	143	48.00	1.69	91.15	88.71	86.99	99.28
Infant during Influenza period	137	57.02	3.54	96.51	93.76	93.76	95.52
Both in utero and infant of during Influenza period	75	47.30	6.49	95.94	95.94	94.45	96.00
In utero during Influenza period	101	49.48	2.22	93.76	92.65	91.56	95.88
Conceived after end of Influenza period	567	52.70	2.89	96.51	94.38	91.93	96.34
Total	1669	51.82	2.41	93.04	90.03	88.18	96.60

Notes: Column (1) refers to sample used in analysis. Column (2) refers to all births observed in SEDD. Column (3) refers to singleton births. Column (7) is restricted to married males.

Table 2: Variable means

		Origin SES		
	Full sample	Skilled	Unskilled	Unknown
Born in parish	19.55	26.58	15.02	8.91
Migrant	80.45	73.42	84.98	91.09
Hög/Kävlinge	58.57	53.66	62.61	58.32
Halmstad/Kågeröd/Sireköpinge	41.43	46.34	37.39	41.68
Age	36.43	36.79	36.42	33.94
Unemployment	3.28	3.21	3.24	4.18
IMR, year of birth+1	62.9	64.96	61.91	56.85
IMR deviation>10%, year of birth+1	25.44	27.69	24.09	21.1
Origin SES				
Skilled	42.3	100	0	0
Unskilled	51.8	0	100	0
Unknown	5.9	0	0	100
Observations	12933	5471	6699	763

Source: Scanian Economic-Demographic Database (SEDD), Bengtsson, Dribe and Svensson 2012.

Table 3: Regression estimates, OLS regression, Individual random effects.

	OLS Regression Estimates, Random Effect at the Individual Level											
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
≥ 5 years of age during Influenza period	-0.382	***	-0.389	***	-0.352	***	-0.372	***	-0.187	***	-0.230	***
	-0.433	*-0.331*	*-0.440*	*-0.339*	*-0.442*	*-0.262*	*-0.425*	*-0.320*	*-0.275*	*-0.099*	*-0.282*	*-0.177*
Between 2 and 5 years of age during Influenza period	-0.321	***	-0.322	***	-0.295	***	-0.317	***	-0.168	***	-0.208	***
	-0.365	*-0.278*	*-0.365*	*-0.278*	*-0.364*	*-0.227*	*-0.361*	*-0.274*	*-0.235*	*-0.101*	*-0.250*	*-0.165*
Between 1 and 2 years of age during Influenza period	-0.200	***	-0.200	***	-0.195	***	-0.212	***	-0.120	***	-0.141	***
	-0.262	*-0.139*	*-0.261*	*-0.138*	*-0.257*	*-0.133*	*-0.274*	*-0.149*	*-0.180*	*-0.060*	*-0.202*	*-0.081*
Infant during Influenza period	-0.155	***	-0.160	***	-0.157	***	-0.172	***	-0.096	***	-0.116	***
	-0.218	*-0.093*	*-0.222*	*-0.098*	*-0.220*	*-0.094*	*-0.235*	*-0.109*	*-0.157*	*-0.036*	*-0.177*	*-0.055*
Both in utero and infant of during Influenza period	-0.108	**	-0.103	**	-0.102	**	-0.115	**	-0.054		-0.070	
	-0.190	*-0.026*	*-0.185*	*-0.022*	*-0.184*	*-0.021*	*-0.197*	*-0.033*	*-0.133*	*0.025*	*-0.149*	*0.010*
In utero during Influenza period	-0.120	***	-0.118	***	-0.123	***	-0.130	***	-0.081	*	-0.088	**
	-0.192	*-0.048*	*-0.190*	*-0.047*	*-0.195*	*-0.050*	*-0.203*	*-0.058*	*-0.151*	*-0.012*	*-0.158*	*-0.018*
Conceived after end of Influenza period	ref		ref		ref		ref		ref		ref	
Age	0.150	***	0.150	***	0.150	***	0.150	***	0.107	***	0.107	***
Age, squared	-0.001	***	-0.001	***	-0.001	***	-0.001	***	-0.001	***	-0.001	***
Origin SES Skilled			ref		ref		ref		ref		ref	
Origin SES Unknown			-0.084	**	-0.084	**	-0.084	***	-0.094	***	-0.094	***
Origin SES Unskilled			-0.108	***	-0.108	***	-0.109	***	-0.109	***	-0.109	***
Migrant			ref		ref		ref		ref		ref	
Born in parish			-0.020		-0.020		-0.019		0.010		0.010	
IMR, year after birth					-0.001				-0.001			
IMR deviation from trend>10%, year after birth							-0.048	*			-0.056	**
Halmstad/Kågeröd/Sireköpinge									ref		ref	
Hög/Kävlinge									0.208	***	0.209	***
Unemployment rate									-0.034	***	-0.034	***
Constant	5.349	***	5.419	***	5.446	***	5.431	***	6.320	***	6.287	***
Observations	12933		12933		12933		12933		12933		12933	
Individuals	1669		1669		1669		1669		1669		1669	
R2, within	0.56		0.56		0.56		0.56		0.58		0.58	
R2, between	0.64		0.65		0.65		0.65		0.67		0.67	
R2, overall	0.51		0.52		0.52		0.52		0.55		0.55	

Note: 90% confidence intervals displayed below parameter estimates for birth cohorts