Germs in the Family: The Short- and Long-Term Consequences of Intra-Household Disease Spread*

N. Meltem Daysal† Hui Ding‡

Maya Rossin-Slater§ Hannes Schwandt¶

March 21, 2023

Abstract

While regular exposure to infectious disease is inevitable for most preschool-aged children, their sickness might exert important externalities on more vulnerable family members, such as their infant siblings. We use Danish population-level administrative data on 35 birth cohorts of children to document a striking difference in the likelihood of severe respiratory illness by birth order: younger siblings have two to three times higher rates of hospitalization for respiratory conditions before age one than older siblings at the same age. The hospitalization gap is larger if the younger sibling is born during seasons of high respiratory disease spread and for siblings with shorter birth spacing, who are prone to close contact. These patterns suggest that the family unit is central in virus transmission, with older children "bringing home" viruses to their younger siblings. We then combine the birth order variation with within-municipality variation in respiratory disease prevalence among preschool-aged children to identify differential long-term impacts of early-life respiratory illness between younger and older siblings. We find that moving from the 25th to the 75th percentile in the local disease prevalence distribution is associated with a 32.4 percent differential increase in the number of respiratory illness hospitalizations in the first year of life for younger compared to older siblings. In the long term, for younger relative to older siblings, we find reductions in educational attainment and earnings at ages 25–32. Lastly, we find a 9.6 percent differential increase in the likelihood of having at least one annual hospitalization for mental health-related causes during adolescence and young adulthood for younger relative to older siblings.

*We thank Carolina Artega, Sarah Cattan, Jon Guryan, Kjell Salvanes, and conference and seminar participants at the Barcelona Graduate School of Economics Summer Forum, the 2021 American Society of Health Economists annual meeting, the 2021 European Economic Association meeting, the 77th Annual Congress of the International Institute of Public Finance, the Ca'Foscari University of Venice, University of Milan, Universita Cattolica, UC Irvine, Michigan State University, Dartmouth College, VATT Institute for Economic Research Helsinki, 15th Nordic Summer Institute in Labor Economics, the Arne Ryde Workshop on Gender and Family Wellbeing, UC Davis, University of Pittsburgh, and Columbia University for helpful comments.

†Department of Economics, University of Copenhagen; CEBI; CESifo; IZA. The activities of CEBI are financed by the Danish National Research Foundation, Grant DNRF134. Email: meltem.daysal@econ.ku.dk.

‡Center for Health and Wellbineg, Princeton University. Email: hd5220@princeton.edu.

§Department of Health Policy, Stanford University School of Medicine; NBER; IZA. Email: mrossin@stanford.edu.

¶School of Education and Social Policy, Northwestern University; NBER; CEPR; IZA. Email: schwandt@northwestern.edu.

1 Introduction

Children get sick frequently, especially when they are in group childcare settings at young ages, and during the fall and winter seasons when common viruses circulate. While regular exposure to infectious diseases is inevitable and beneficial for training their immune systems (Holt and Jones, 2000; M'Rabet et al., 2008; Côté et al., 2010; Van den Berg and Siflinger, 2020; Fink et al., 2021), preschoolers' sickness might exert important externalities within their families, especially on younger infant siblings who are in a vulnerable period of rapid brain development (Eppig et al., 2010; Bhalotra and Venkataramani, 2013). Yet despite the universality of this experience among families with young children, there is limited empirical evidence quantifying such within-family externalities.

This paper focuses on the spread of respiratory illnesses among young children and studies the magnitudes of these externalities over the short and long-run. We use population-level Danish administrative data covering 35 birth cohorts to study: (i) how respiratory illnesses spread from older to younger siblings during their first year of life, when they are particularly vulnerable to severe disease and complications, and (ii) how respiratory disease exposure during infancy affects the younger siblings' long-term health, human capital, and economic outcomes.

We begin by documenting a striking disparity in the likelihood of severe respiratory disease in early childhood by birth order. Using data on all first- and second-born siblings born in Denmark between 1981 and 2015, we find that younger siblings have two to three times higher rates of hospitalization for respiratory conditions during their first year of life compared to the older siblings at the same age, and that this gap is particularly large when hospitalizations are measured in the first three months of life.[1] Moreover, the hospitalization disparity is larger if the younger sibling is born in the fall or winter, when respiratory viruses circulate more frequently. The hospitalization gap is also larger for siblings with shorter birth spacing, who may be more prone to close contact that facilitates virus transmission. These patterns highlight the family unit as being central in virus transmission, and the hitherto under-studied

[1] Note that this finding builds on several prior studies that document that higher-order siblings have *better* health outcomes at birth than first-borns (e.g., Brenøe and Molitor, 2018; Pruckner et al., 2021). Thus, it appears that younger siblings are more susceptible to severe respiratory infection, despite having better health at birth than their older counterparts.

mechanism by which birth order might influence children's longer-term outcomes—older children "bring home" common viruses (e.g., from group childcare environments), putting their younger siblings at heightened risk of severe respiratory illness in the first few months of life.

While younger siblings are more likely to experience severe respiratory illness in the first few months of life than their older counterparts, the long-term impacts of this differential likelihood of illness are ambiguous. On the one hand, the expansive literature on a wide range of adverse shocks in early childhood documents lasting damages to human capital formation and other measures of adult well-being (Currie and Almond, 2011; Almond et al., 2018). On the other hand, evolutionary biology studies highlight the importance of physiological adaptation—i.e., that adverse shocks can lead to beneficial biological adaptations in humans (Bateson et al., 2014; Gluckman and Hanson, 2006)—and identify a high rate of immune system learning in the first year of life (Holt and Jones, 2000; M'Rabet et al., 2008; Côté et al., 2010). Thus, exposure to an infectious disease in infancy may increase immunity for an individual if they are exposed to the same virus at older ages, suggesting a potentially non-linear relationship between early-life exposure and long-term outcomes (Van den Berg and Siflinger, 2020; Fink et al., 2021). This type of immunity formation is particularly important for understanding the impacts of *endemic* viruses to which children are exposed on a regular basis throughout their lives.

To identify the long-term causal impacts of early-life respiratory disease exposure, we combine the birth order variation in the likelihood of severe respiratory infection together with variation in local disease prevalence. Local respiratory disease prevalence among children is largely driven by highly infectious conditions, such as the Respiratory Syncytial Virus (RSV), which spread across locations in irregular waves (Pitzer et al., 2015).[2] We construct a municipality-level index, which is designed to capture respiratory disease exposure during each child's first year of life from slightly older children in the community. We calculate the number of hospitalizations for respiratory conditions per 100 children aged 13 to 71 months in each municipality, and then assign to each child the cumulative child hospitalization rate

[2] As demonstrated by Pitzer et al. (2015), climatic factors—including temperature, vapor pressure, precipitation, and potential evapotranspiration—are important predictors of geographic variation in RSV transmission rates. While these factors may have impacts on long-term outcomes through channels unrelated to respiratory disease spread (see, e.g., Isen et al., 2017a, for evidence on early-life exposure to extreme temperature), we note that such channels are unlikely to differentially influence first versus second-born children.

in their municipality over their first 12 months of life.[3] We then use our sample of siblings to estimate the differential effect of the disease index for younger compared to older siblings. Our regressions control for birth order, municipality, birth year, and birth month fixed effects, thus accounting for other differences between older and younger siblings, time-invariant differences across municipalities that might drive differences in disease exposure, and aggregate and seasonal trends in respiratory illness.

We show that the local respiratory disease index strongly predicts the likelihood that a child is hospitalized for a respiratory illness during the first year of life, and that this impact is much larger for younger relative to older siblings. We find that moving from the 25th to the 75th percentile in the disease index distribution is associated with a 0.022 differential increase in the number of respiratory illness hospitalizations in the first year of life for younger relative to older children, representing an additional 32.4 percent increase at the sample mean. This effect is in part driven by a differential increase in hospitalizations for RSV, which is a mild illness in most older children but can be extremely serious among infants.[4]

In the long run, increased exposure to severe respiratory illness during infancy among second-born children impacts their health, human capital, and labor market outcomes. We find that, for the younger siblings, moving from the 25th to the 75th percentile in the disease index distribution exposure in the first year of life is associated with a 0.01 of a standard deviation penalty in ninth grade Danish test scores. We also find some evidence of a delay in high school graduation—younger siblings are 0.5 percentage points less likely to graduate by age 20, although there is some indication that they may catch up at older ages. We see similarly negative impacts on college graduation around ages 23–25. When it comes to adult labor market outcomes, we find that moving from the 25th to the 75th percentile in the disease index distribution leads to a 0.7 percent reduction in income conditional on employment at ages 25–32 and a 0.3 percentage point decline in income percentile rank at the same ages

[3] If a given child has an older sibling who is between 13 and 71 months of age during their first year of life, we exclude the older sibling from the hospitalization rate.

[4] In most healthy individuals, RSV causes mild, cold-like symptoms. But in infants, RSV can cause severe respiratory infections, including bronchiolitis and pneumonia. Recent estimates suggest that approximately 14.7 per 1,000 infants under six months of age and 2.9 per 1,000 children under age five are hospitalized with RSV every year (Rha et al., 2020). For comparison, the highest COVID-19 hospitalization rate for children ages 0–4 so far (in January 2022, at the height of the Omicron wave) is estimated to be 0.16 per 1,000 (see: `https://gis.cdc.gov/grasp/covidnet/covid19_3.html`).

among younger relative to older siblings (and we do not find any significant impacts on the extensive margin of labor force participation). The magnitudes of our estimated effects on long-term earnings are comparable to the effects of an 8 percent reduction in birth weight (Black et al., 2007) or a 7 percent increase in ambient air pollution in one's year of birth (Isen et al., 2017b); they are smaller than the impacts of *in utero* exposure to the 1918 Spanish Influenza pandemic (Almond, 2006) or *in utero* exposure to a maternal influenza infection that requires hospitalization (Schwandt, 2018).

We additionally examine the impacts of respiratory illness exposure in the first year of life on hospitalizations for respiratory conditions in later childhood. We find that higher respiratory disease exposure before age one is associated with a *lower* likelihood of hospitalization for all respiratory conditions at ages three to four, consistent with the hypothesis of immunity formation.[5] This protective effect disappears after age four, when infancy disease exposure stops being associated with hospitalizations for respiratory diseases. Thus, while the protective effects of infancy exposure to respiratory diseases are limited to the first few years of childhood, there also does not appear to be an adverse impact on respiratory health in later childhood or early adulthood.

A likely biological mechanism for the long-term effects on human capital and labor market outcomes that we find is the impairment of brain development during infancy (Adams-Chapman and Stoll, 2006; Bilbo and Schwarz, 2012; O'Shea et al., 2013). As summarized by Bhalotra and Venkataramani (2013), the biomedical literature emphasizes the importance of fast neural development during that period coupled with a high degree of neural plasticity. During infancy, about 85 percent of calorie intake is used for neural growth (Eppig et al., 2010), and severe illness can both reduce calorie intake as well as divert calories away from brain development to fighting the disease. Deverman and Patterson (2012) argue that inflam-

[5] At the same time, we do not observe a protective effect on the likelihood of subsequent hospitalization for RSV. This result is consistent with RSV being only a partially immunizing disease—that is, an RSV infection does not provide full immunity against future illness (Lambert et al., 2014; Fuentes et al., 2016). This lack of immunity formation, combined with the fact that RSV accounts for a large share of all respiratory hospitalizations during infancy (30 percent among second-born children), suggests that RSV might be a particularly important driver of the adverse long-term impacts on educational and economic outcomes. Unfortunately, we cannot measure the long-term effects of RSV illness directly, as RSV is not contained in the International Classification of Disease version 8 (ICD-8) coding system that was used in Denmark until 1994. Thus, we can only measure RSV exposure for cohorts born in 1994 and later, when Denmark switched to the ICD-10 system, and these cohorts are too young to measure adult outcomes through age 32 in our data.

matory responses to illness can also directly impair brain development.[6] These illness-driven disruptions of brain development in infancy can go on to impact later-life cognitive and mental health outcomes, both of which are important inputs into human capital attainment and economic productivity (see, e.g., Bütikofer et al., 2020; Biasi et al., 2021).

To further investigate the mechanisms driving the estimated long-term impacts on adult human capital and economic outcomes, we study mental health care utilization in adolescence and young adulthood. We show that moving from the 25th to the 75th percentile in the respiratory disease index distribution during the first year of life is associated with a 0.04 percentage point (9.6 percent) increase in the likelihood of experiencing any hospitalizations with a (primary or non-primary) mental health-related diagnosis at ages 16–26, and a 0.06 percentage point (5.0 percent) increase in the likelihood of having any visits to psychiatrists at the same ages. These mental health impacts are within the range of estimates of the impacts of fetal and early childhood shocks on later mental health outcomes in the existing literature, including exposure to Ramadan (Almond and Mazumder, 2011), maternal stress (Persson and Rossin-Slater, 2018), and changes in economic conditions (Adhvaryu et al., 2019).

Lastly, we analyze heterogeneous impacts on both respiratory hospitalizations during infancy and long-run outcomes along a variety of dimensions, including parental socio-economic status, the younger sibling's gender and health at birth, child birth spacing, and whether the older sibling is in a childcare center. When it comes to the short-run effects on respiratory hospitalizations, we find that the effects are disproportionately concentrated among low birth weight younger siblings (those with birth weight less than 2,500 grams). Further, younger male siblings experience a larger differential increase in respiratory hospitalizations than their female counterparts, which is consistent with the "fragile male" hypothesis (i.e., the idea that male fetuses and infants are biologically more vulnerable to various shocks and stressors, see, e.g. McCarthy, 2019; Sanders and Stoecker, 2015; Kraemer, 2000). The effect on hospitalizations also seems to be monotonically decreasing with birth spacing—that is, younger siblings in families with a shorter birth spacing gap have a larger differential increase in hospitalizations before age one. The estimated impact on respiratory hospitalizations is also larger in

[6] Medical treatment occurring during hospitalization for severe respiratory illness has the potential to additionally harm brain development, e.g., when infants are put into medically induced coma to allow for prolonged ventilation (Vliegenthart et al., 2017).

sibling pairs in which the older child is attending a childcare center compared to pairs in which the older child is not. These patterns support the conjecture that intra-family spread is a key mechanism in driving higher rates of respiratory illness among younger siblings.[7]

When it comes to heterogeneity in long-run impacts, we do not find evidence of significant differences in effects on test scores or educational attainment. The adverse effects on adult income appear to be larger among males than females, consistent with what we see for respiratory hospitalizations during the first year of life. At the same time, we find that the effects on mental health care utilization in adolescence and young adulthood are concentrated among females. One potential explanation for this pattern could be that while respiratory disease exposure in infancy affects underlying mental health among both genders, young women may be more likely to seek care and treatment than young men (Pattyn et al., 2015). This disproportionate use of mental health care might in turn buffer against the adverse impacts on later economic productivity among women more than men.[8]

This study contributes to an expansive body of work on the human capital impacts of early life circumstances (Barker, 1990; Currie and Almond, 2011; Black et al., 2017; Almond et al., 2018). This literature includes estimates of the impacts of a vast range of prenatal and early childhood factors—from economic resources (e.g., Hoynes et al., 2016; Adhvaryu et al., 2019; Bailey et al., 2020) to nutrition (e.g., Almond and Mazumder, 2011) to environmental conditions (e.g., Almond et al., 2009; Isen et al., 2017b; Black et al., 2019) to maternal stress (e.g., Black et al., 2016; Persson and Rossin-Slater, 2018). The literature on infectious diseases in early childhood has focused on severe infectious diseases, such as malaria, measles, and polio, that have been largely eliminated in high-income countries but still exist in the developing world (Bleakley, 2010; Barreca, 2010; Cutler et al., 2010; Lucas, 2010; Venkataramani, 2012; Chang et al., 2014; Barofsky et al., 2015; Gensowski et al., 2019; Kuecken et al., 2021; Fink et al., 2021; Chuard et al., 2022), and on large-scale pandemics like the 1918 Spanish Flu

[7] Further, the heterogeneous results by birth spacing suggest that our effects are *not* driven by differences in parental investments between older and younger siblings (and the potential interactions between these investments and our disease indices). Price (2008) finds that in the U.S., the difference in parent-child quality time between first- and second-born children is larger when the birth spacing gap is longer. Our pattern of a monotonically decreasing effect with birth order is the opposite of what would be predicted if differential parental time investment were the main channel.

[8] Given that we do not have any way to observe underlying (untreated) mental illness in our data, however, we unfortunately cannot provide any additional empirical support for this conjecture.

(Almond, 2006; Almond and Mazumder, 2005; Lin and Liu, 2014) and the 1957 Asian Flu (Kelly, 2011). Schwandt (2018)'s analysis is an exception in that it focuses on the impacts of exposure to an endemic respiratory virus—the seasonal influenza—but only during the *in utero* period. Our study builds on this work by studying a range of respiratory illnesses that circulate among young children on a regular basis, and by focusing on the first year of life instead of the prenatal stage.[9] Our novel estimates of long-term impacts of severe respiratory disease can inform household behaviors and cost-benefit evaluations of policies designed to curb transmission of common viruses, including vaccination mandates, drug distribution programs, and sick pay regulations (Bhalotra and Venkataramani, 2015; White, 2019; Pichler and Ziebarth, 2020; Bütikofer and Salvanes, 2020; Atwood, 2022; van den Berg et al., 2023).

Our analysis further contributes to the literature on birth order and sibling spillovers, which has documented worse human capital and life outcomes for later-born children relative to first-borns (Black et al., 2005; De Haan, 2010; Buckles and Kolka, 2014; Brenøe and Molitor, 2018; Lehmann et al., 2018; Breining et al., 2020; Black et al., 2021). This literature typically points to family resources and uneven parental investments as drivers of younger siblings' disadvantage (Price, 2008). Our results suggest that the disease environment during infancy is an additional source of disadvantage for later-born children, and that the older sibling likely serves as a vector of transmission. Importantly, the long-term effects we measure are net of any parental responses to the health shocks. To the extent that parents may respond to one child's sickness in a compensatory way—as found by Yi et al. (2015) and Daysal et al. (2020)—the sibling differences in long-run outcomes that we find represent lower bound estimates of the uncompensated (i.e., "biological") impacts of respiratory illness during infancy on later well-being.

Finally, this study is also relevant for the assessment of the costs of the COVID-19 pandemic for young children. While children have been considered to be a low-risk group for infection

[9]Studies in the medical literature have analyzed the health impacts of RSV infection, with a focus on asthma as an outcome. These studies use relatively small samples of children to correlate RSV infection (or RSV hospitalization) with later health conditions (e.g., Kneyber et al., 2000; Korppi et al., 2004; Kusel et al., 2007; Régnier and Huels, 2013; Zomer-Kooijker et al., 2014; Carbonell-Estrany et al., 2015). Related, a recent study using data from Finland reports on the association between being hospitalized for any infection at ages 0–18 and adult economic outcomes (Viinikainen et al., 2020). We are not aware of studies using quasi-experimental designs to isolate causal impacts of early life RSV exposure, or those using population-level administrative data.

with the SARS-CoV-2 virus, the pandemic may have lasting and dynamic impacts on children through its effects on other infectious diseases. Policies implemented during the pandemic—including travel restrictions and school closures—have reduced the spread of other respiratory viruses, such as RSV (Leung et al., 2020; Cowling et al., 2020). At the same time, the spread of RSV and other common respiratory viruses surged in 2021 and 2022 once the restrictions were lifted, reflecting a larger than usual susceptible population of young children who had been shielded during the early stages of the pandemic. Our results suggest that infants with older siblings may have benefited from the pandemic-induced muted disease spread during the first year of the pandemic, while those born during the following two years might have experienced stronger than usual disease exposure. Thus, the COVID-19 pandemic may have differential long-term effects on children born before and during the pandemic through its dynamic impacts on the spread of other infectious diseases that are more serious in early life than COVID itself, including RSV.

2 Data and Sample

We use several population-level administrative data sets from Denmark in our analysis. These data include individual-level records with unique personal identifiers that allow us to follow individuals over time and to link family members to one another.

Outcomes. Our key short-run outcome is the number of hospitalizations with a primary diagnosis of a respiratory illness during the first year of life. We measure this outcome using the *National Patient Register*, which is available to us for years 1981–2016 and includes all inpatient admissions to public and private hospitals, along with International Classification of Disease (ICD) diagnosis and procedure codes (Lynge et al., 2011). Denmark used the International Classification of Disease version 8 (ICD-8) coding system until 1994, and then switched to the ICD-10 system for all years going forward.

We classify inpatient visits with the following primary diagnosis codes as respiratory disease-related: ICD-8 codes starting with "46," "47," "48," "490," "079," and "783"; and ICD-10 codes starting with "B974" or "J" (excluding "J4"). In additional models, we examine hospitalizations for RSV, using data on cohorts born in 1994 and later, which we identify with

ICD-10 codes J12.1 (respiratory syncytial virus pneumonia), J20.5 (acute bronchitis due to respiratory syncytial virus), J21.0 (acute bronchiolitis due to respiratory syncytial virus), and B97.4 (respiratory syncytial virus as the cause of diseases classified elsewhere).[10]

To study human capital outcomes in later life, we consider ninth grade Danish (reading) and mathematics test scores from the *Academic Achievement Register* for years 2001–2019, which we standardize within subject and test year such that they have a mean of zero and a standard deviation of one. We also use information on the highest level of completed schooling, which comes from the *Education Register* for years 1981–2019, and is drawn from administrative school records. We study two long-run educational outcomes: indicators for having graduated from high school and from college, respectively, measured by ages 18 through 32.

We use two registers to measure labor market outcomes. We use the *Register-Based Labour Force Statistics* available for years 1980–2019 to characterize labor force participation. This data set is based on tax records, and records the labor market status of the entire Danish population (observed on January 1st) as of November of the preceding year (Petersson et al., 2011). We construct an indicator equal to one if an individual is in the labor force and zero otherwise (i.e., those who are employed and unemployed but searching are both coded as 1; those out of the labor force are coded as 0). We also use the *Income Statistics Register* for years 1980-2019 to calculate the natural log of gross personal income, converted into 2010 $USD. Lastly, we create a variable that denotes the percentile rank of an individual's gross personal income in the overall Danish population (i.e., not just our sample) in each birth cohort and at each observed age. We study these labor market outcomes at ages 18 through 32.

Finally, we consider two outcomes capturing mental health care utilization at ages 16 through 26. First, we study mental health-related hospitalizations using inpatient admissions that have either a primary or non-primary diagnosis code starting with "29," "30," or "31" in ICD-8 format or "F" in ICD-10 format. Second, we use the *Health Insurance Register* to study visits to psychiatrists. These data, available to us for years 1997–2015, provide information on reimbursements to private-practice physicians – both general practitioners and specialists – for all patient-related services covered by the national health insurance. We identify psychiatrist

[10] We can only measure RSV from 1994 onward (when ICD-10 was used in Denmark) because the ICD-8 system did not have any codes specific to RSV.

visits based on the physician's specialty code ("24" or "26").

Control variables. We observe a rich set of child and parent characteristics, using the previously described registers as well as the *Population Register* and the *Birth Register*. The *Population Register* provides a snapshot of demographics on all Danish residents as of January 1st of each year (Pedersen, 2011). The *Birth Register* includes the universe of births, with information on the exact date of birth, gender, plurality, and birth weight. It also has unique parental identifiers, allowing us to link siblings and determine birth order.[11]

We include the following variables as controls, measured at the time of childbirth: child gender, maternal age, maternal foreign-born status, maternal education level, and parental marital/cohabitation status.[12] We also include controls for the natural log of the mother's, father's, and the family's total income, as well as each parent's employment status, all measured in the year before childbirth. Lastly, we include the birth spacing between siblings in months, as well as an indicator for being the younger sibling interacted with birth spacing.

Finally, in some of our heterogeneity analyses, we make use of a data set containing information on children's enrollment in Danish childcare centers, which is reported annually in September of each year. This information is available to us over the period of September 1995 to September 2013.

Analysis sample. To construct our analysis sample, we begin with the universe of 2,221,433 children born between 1981 and 2015 in Denmark and make the following restrictions. First, we exclude families with only one child. Second, we only keep the first and second-born children in every family, and further, we only keep families in which the first and second-born children are singletons. Third, we only keep children in sibling pairs with a birth spacing gap of at least 11 months, which ensures that there is no overlap in the first year of life of the two children. Fourth, we only keep children with non-missing information on municipality of birth and who are born in municipalities that have an average of at least 1,000 children aged

[11]Specifically, the birth records contain identifiers for all mothers. If the mother is married at the time of childbirth, then her husband is automatically registered as the biological father. If the mother is unmarried, then the biological father's identifier is listed if he establishes paternity. Fathers' identifiers are missing for only 0.58% percent of observations in our analysis period.

[12]Information on parental marital/cohabitation status is collected from the year after birth, due to the lag in administrative record.

13–71 months over the sample period, which ensures that we have sufficient observations to calculate the respiratory disease exposure index as described in Section 3 below.[13] Finally, we drop children with missing parental control variables, and keep sibling pairs in which both children remain in the sample after these restrictions. Appendix Table A1 shows how our sample size evolves as we make these various restrictions to arrive at our final analysis sample.

Our final analysis sample consists of 1,163,982 children, which we use to analyze short-term impacts of respiratory disease exposure on hospitalizations in the first year of life. When studying long-term outcomes, our sample sizes differ depending on the ages at which outcomes are measured. To study test scores, we use children born between 1986 and 2003 because test score data begin in 2001 and we need to observe children when they are in ninth grade (around age 16). To study mental health care at ages 16–26 and educational attainment and labor market outcomes at ages 18–32, we analyze children born in cohorts who can be observed in our outcome data at those ages.

3 Descriptive Analysis and Empirical Design

3.1 Differences in Respiratory Disease Hospitalizations between Older and Younger Siblings

We begin with a descriptive analysis of respiratory disease hospitalization patterns among children in our sample, comparing first- and second-born siblings. This analysis sheds light on a likely mechanism through which respiratory diseases spread within families—older children, most of whom interact with same-age peers in group childcare settings and are therefore frequently exposed to infectious viruses, "bring home" diseases that infect their younger siblings.

Raw sibling differences. Panel (a) of Figure 1 plots the average number of respiratory disease hospitalizations (per 100 children) by child age in months during the first year of

[13]Denmark changed its administrative municipality structure in 2007, which led to a reduction in the total number of municipalities from 275 to 98. We use the current municipality structure in our analysis, and use a crosswalk that matches each pre-2007 municipality to the appropriate municipality code used from 2007 onward. When dropping municipalities with an average of fewer than 1,000 children aged 13–71 months over the sample period, we drop 7 municipalities, such that our final analysis sample contains 91 municipalities in total.

life. It shows that, compared to first-born children, younger siblings have two to three times higher rates of hospitalization for respiratory disease, and that the difference is especially large when children are two and three months of age. Panel (b) of Figure 1 extends the time horizon on the $x-$axis to 60 months (i.e., age five), and demonstrates that the difference in hospitalization rates between older and younger siblings disappears after age one. This pattern is consistent with the vast majority of Danish children staying home with their mothers during their first year of life, and only starting to attend group childcare after they turn one year old.[14] Thus, after age one, younger and older siblings are similarly likely to be exposed to infectious viruses in group care environments, whereas non-first-borns have exposure before they turn one through their older siblings bringing viruses home.[15]

Seasonal differences. In Figure 2, we explore the role of respiratory disease seasonality in driving the observed hospitalization gap between siblings. The two graphs in Figure 2 show the average number of respiratory disease hospitalizations for older and younger siblings, respectively, separately by season of birth. These graphs reveal three facts. First, children are more likely to be hospitalized for respiratory disease during the winter when common respiratory disease outbreaks (such as RSV) are more prevalent—children born in November, December, and January have the highest hospitalization rates in the first three months of life; those born in August, September, and October have the highest hospitalization rates at 3 to 6 months old; those born in May, June, and July have the highest hospitalization rates at 7 to 9 months old; and those born in February, March, and April have the highest hospitalization rates at 10 to 12 months old. Second, younger siblings have higher hospitalization rates than older siblings regardless of season of birth. Third, out of all sub-groups considered, younger siblings born in the winter months have the highest hospitalization rates when they are two

[14]In Denmark, some form of maternity leave has been available since the beginning of the 20th century. In 1980, mothers had access to 14 weeks of nearly fully paid leave following the birth of a child, and this leave benefit was extended to 24 weeks (and also began to include fathers) in 1985 (Rasmussen, 2010). Subsequently, additional weeks of leave were added with reduced benefit compensation. By 2002, new parents could receive up to 52 weeks of parental leave with partial pay. The majority of this leave is used by mothers (see, e.g., Beuchert et al., 2016).

[15]Appendix Figure A1 plots the share of children enrolled in a group childcare center, nursery, or preschool by age in months. Virtually no children attend childcare before they turn one year old, and the share increases rapidly over ages one to two. There is a small jump at age three, when children are eligible to attend formal preschool centers (as opposed to less formal nurseries for younger children). More than three-quarters of children are enrolled in a center by the time they are three years old.

to three months old, suggesting that they are particularly susceptible to severe respiratory infections during early infancy.

Birth spacing differences. Lastly, in Figure 3, we examine differences in these patterns across siblings with different birth spacing gaps. Each graph plots the average number of respiratory disease hospitalizations per 100 children by age in months of the older siblings (on the left) and the younger siblings (on the right), separately by season of birth and for different birth spacing gaps. The graphs demonstrate that younger siblings born in winter months have the highest hospitalization rates regardless of birth spacing, and that the difference in hospitalizations between younger and older siblings gets much smaller as birth spacing increases. This pattern is consistent with siblings having more interactions that facilitate disease spread when their age difference is smaller, and with the older siblings—i.e., the ones who "bring home" disease—being more susceptible to infection when they are younger themselves (since the age of the older siblings observed in the right-hand graphs in Figure 3 falls when the birth spacing gap is smaller).

In sum, the observed patterns in the data—(i) higher hospitalization rates among younger siblings than older siblings, (ii) a larger sibling hospitalization gap during the winter season, and (iii) a larger hospitalization gap for more closely spaced siblings—are consistent with the idea that respiratory disease spreads within the family because older children "bring home" viruses that they pick up in their local community (e.g., at their childcare center). This analysis informs our empirical strategy for estimating the causal effects of early childhood respiratory disease exposure: We focus on exposure during the first year of life, leverage variation in local respiratory disease outbreaks among slightly older children, and analyze differential effects across older versus younger siblings.

3.2 Empirical Strategy for Estimating Causal Effects of Early Life Respiratory Disease Exposure

Our main independent variable is designed to capture respiratory disease exposure during the first year of life from slightly older children in the local community. We begin by using the *National Patient Register* data to obtain the number of respiratory disease hospitalizations per 100 children aged 13 to 71 months in each municipality and calendar year-month over our analysis time frame.[16] To allow for an informative visualization of the variation in this respiratory hospitalization rate, in Appendix Figure A2, we plot the raw month-by-month values of the rate in each of Denmark's 10 most populated municipalities, separately over four time periods during our sample time frame: 1980–1989, 1990–1999, 2000–2009, and 2010–2016. Consistent with our descriptive analysis above, we observe a strong seasonal pattern, with a higher hospitalization rate during the winter months in all locations and across all time periods. At the same time, there is a substantial amount of variation in children's respiratory hospitalizations across municipalities in any given month, as well as within each municipality over time. In Appendix Figure A3, we demonstrate the central source of variation used to identify the key estimates in our empirical model (described in more detail below)—we use data for all municipalities in Denmark for the entire sample period, regress the hospitalization rate on municipality, year, and month fixed effects, and plot the distribution of the residuals. The figure demonstrates that there remains a substantial amount of variation in respiratory disease hospitalizations even after location and time fixed effects are partialled out.

Next, for each child in our sibling analysis sample, we assign this monthly respiratory hospitalization rate to each month of their first year of life based on their municipality of residence in that month. Importantly, if a given child has an older sibling who is between 13 and 71 months of age at any point during their first year of life, we exclude the older sibling from the hospitalization rate. Finally, we define the disease exposure index as the sum of the monthly hospitalization rates over the 12 months of each child's first year of life. Thus, our index captures a child's cumulative respiratory disease exposure before age one from slightly older children in their municipality.

Our empirical models estimate the differential effect of the respiratory disease exposure index on younger versus older siblings. Specifically, our regression models take the form:

$$Y_{itkm} = \beta_0 + \beta_1 Younger_i + \beta_2 Index_{itkm} + \beta_3 Younger_i \times Index_{itkm} + \mu_m + \theta_t + \rho_k + \gamma' X_i + \epsilon_{itkm} \quad (1)$$

for each child i born in year t, month k, and municipality m. Y_{itkm} is an outcome such as the

[16] We use 71 months (i.e., 5 years and 11 months) as the upper age limit to capture respiratory disease spread among preschool-aged children, most of whom are in group childcare environments. Children start primary school at age 6 in Denmark.

number of hospitalizations during the first year of a child's life that have a primary diagnosis of a respiratory condition, or an indicator for having graduated high school by age 20. $Younger_i$ is an indicator set to 1 for younger siblings, and captures the "main" effects of birth order on our outcomes of interest. $Index_{itkm}$ is the respiratory disease exposure index described above. μ_m are municipality fixed effects that account for time-invariant geographic differences in exposure to infectious diseases and in other determinants of our outcomes. θ_t and ρ_k are year and month of birth fixed effects, respectively, that control for cohort and seasonal trends. X_i is a vector of individual and family background control variables measured in the year of birth: indicator for the child being male, the birth spacing between siblings in months and birth spacing interacted with the younger siblings indicator, mother's age and age squared, indicator for mother's foreign-born status, indicators for mother's education level (high school degree, college degree or higher), and an indicator for parents being married or cohabiting. We also control for the natural log of the mother's, father's, and total family income, as well as indicators for each parent being employed, in the year before childbirth. We cluster standard errors at the municipality level.

Identifying assumption. The key coefficient of interest in model (1), β_3, measures the differential impact on younger siblings relative to older siblings of an additional respiratory disease hospitalization per 100 children aged 13–71 months in the child's municipality during their first year of life. Interpreting this coefficient as representing a causal impact of respiratory disease exposure relies on an assumption that there are no unobserved municipality-specific time-varying factors that are (a) correlated with respiratory disease prevalence, (b) influence children's outcomes, and (c) differentially impact younger versus older children in a family. While this assumption is not directly testable, we assess its plausibility in several ways.

First, we investigate the sensitivity of our main results across specifications that include various controls, including municipality-specific linear trends, and mother fixed effects. As we show below, our results are fairly robust across these models.

Second, we estimate model (1) without the controls in X_i and instead using the X_i variables as outcomes (Pei et al., 2019). We additionally consider two other relevant placebo outcomes in this context: indicators for low birth weight (less than 2,500 grams) and very low birth weight (less than 1,500 grams) births. Results are presented in Appendix Table A2. We find that three out of the 14 interaction coefficients reported in this table are statistically significant but very small in magnitude—mothers of younger siblings are slightly older and the parents are slightly less likely to be married or cohabiting in municipalities with higher respiratory disease exposure indices; additionally a higher disease index during the younger child's first year of life is correlated with a slightly longer birth spacing of about 1 month. We control for birth spacing interacted with the younger child indicator, as well as maternal age and parental marital/cohabitation status in all of our analyses.

Third, we construct two alternative indices, in which instead of using children's hospitalizations for respiratory conditions, we use: (i) non-infectious digestive diseases, and (ii) injuries and poisonings. If the differential likelihood of hospitalization for respiratory conditions for younger compared to older children reflects differences in parental healthcare-seeking behavior (i.e., parents are more likely to go to the hospital with their second-born than their first-born at the same level of underlying illness), then we might expect similar patterns to emerge for other *non-infectious* childhood health shocks, such as those stemming from digestive issues or accidents. Yet when we estimate model (1) using the two alternative indices and hospitalizations in the first year of life for these causes, we do not find evidence in support of this hypothesis (see Appendix Tables A3 and A4). If anything, we find that younger children are less likely to be hospitalized for these causes, and there is no evidence of significant positive interactions between the alternative indices and the younger child indicator.[17]

Overall, these analyses support our identifying assumption, and suggest that our model is likely to yield causal estimates of the differential effects of respiratory disease exposure in early childhood for younger relative to older siblings.

Sample means. Table 1 presents means of some of the key variables in our analysis, separately for the older and younger siblings in the sample. The table highlights some important differences in child outcomes by birth order. Compared to older siblings, younger siblings have higher average birth weight (3589 versus 3431 grams for younger versus older siblings, respec-

[17]Note that the significant main effect of the injury index on hospitalizations for the same causes is plausible, as they are likely driven by underlying local and seasonal factors (e.g., icy conditions may increase the local injury rate among children).

tively). The average values of the respiratory disease exposure index for older and younger siblings are similar: 2.8 and 2.9 hospitalizations per 100 children, respectively. However, despite the slight advantage in health at birth (which has been found in other settings, see, e.g. Brenøe and Molitor, 2018; Pruckner et al., 2021) and similar local exposure to respiratory disease, younger siblings' average number of hospitalizations for respiratory conditions during their first year of life is nearly *twice* the average for older siblings (9.0 and 4.7 per 100 children for younger and older siblings, respectively). The relative difference is even larger for RSV hospitalizations during the first year of life, with younger siblings' average number of hospitalizations *three times higher* relative to older siblings.[18] Moreover, consistent with prior literature on the impacts of birth order (e.g., Black et al., 2005), younger siblings have worse educational and economic outcomes than their older counterparts. Additionally, younger siblings have higher rates of mental health care utilization, as measured by hospitalizations for mental health-related conditions and visits to psychiatrists.

The table shows that mothers are on average aged 26.8 years at the time of their first birth and 30.3 years at the time of their second birth. Approximately 4.5 percent of mothers in our sample are foreign-born. About 75.0 and 78.9 percent of mothers have a high school degree at the time of the first and second birth, respectively, while 30.2 and 36.7 percent have a college degree, respectively. Approximately 93.7 percent of parents are married or cohabiting at the time of the first birth, while 95.0 percent are married or cohabiting at the time of the second. Household income is slightly higher at the time of the second than the first birth.

4 Results

In this section, we first discuss our results on the relationship between the respiratory disease index and hospitalizations for respiratory conditions during childhood, for younger versus older siblings. We then discuss our results on long-run educational and labor market outcomes, which we can measure through age 32 in our data. We additionally present our results on mental health care utilization in adolescence and young adulthood. We follow up by bench-

[18] The average number of hospitalizations for all respiratory conditions among the 1994+ cohorts, for whom we observe RSV-specific hospitalizations, is similar to the overall sample that includes older cohorts: 9.9 and 4.5 per 100 children for younger and older siblings, respectively.

marking the magnitudes of our long-run estimates against those found in the prior literature, and provide some additional sensitivity analyses.

4.1 Short-Term Effects of Respiratory Disease Exposure on Respiratory Hospitalizations

Table 2 presents results from estimating equation (1) using as the outcome the number of hospitalizations during the first year of a child's life that have a primary diagnosis of a respiratory condition. We report the coefficients on the indicator denoting the younger sibling, the respiratory disease exposure index (expressed as the number of respiratory disease hospitalizations per 100 children aged 13 to 71 months), and the interaction of these two variables. Column (1) shows that, consistent with the graphical evidence in Figures 1 through 3, younger siblings on average have 0.039 more (57.4 percent relative to the sample mean) hospitalizations for a respiratory condition before age one than their older counterparts. Column (2) shows that there is a positive correlation between the disease exposure index and the likelihood of hospitalization before age one in the overall siblings sample, and column (3) demonstrates that the coefficients on the younger sibling indicator and the disease exposure index do not change when they are both included in the same regression model. Once we include the interaction term in columns (4) and (5), we find that there is a significant differential effect of local respiratory disease exposure on younger siblings relative to older siblings. In particular, we find that an additional respiratory hospitalization per 100 children aged 13–71 months in a municipality increases the younger sibling's number of hospitalizations during the first year of life by an average of 0.012 (17.6 percent), as compared to the older sibling. This relationship is robust across specifications without and with family background control variables (columns 4 and 5, respectively). In the bottom row of the table, we report the magnitude of the differential effect on younger siblings relative to older siblings of an increase in the disease exposure index from the 25th to the 75th percentile of the index distribution. This magnitude amounts to a 0.022 differential increase in the number of respiratory disease hospitalizations in the first year of life, which represents an additional 32.4 percent relative to the sample mean.

In Figure 4, we explore heterogeneity in our estimates. We estimate our baseline model (1), and include subgroup indicators interacted with the younger sibling indicator, the disease index, and the younger sibling indicator × disease index interaction. We then plot the coefficients and 95% confidence intervals from estimates of the triple interaction terms. We consider differences in effects by: parental socio-economic status (defined as the mother's years of education being above or below the median in the distribution), child health at birth (low-birth-weight and non-low-birth-weight children), child gender, the birth spacing between the siblings, and whether or not the older child is enrolled in a childcare center (limited to siblings with a birth spacing gap of no more than 2 years).[19] We find that the effects on respiratory hospitalizations are much larger for younger siblings who are low birth weight than those who are not. Additionally, consistent with the "fragile male" hypothesis regarding the biological vulnerability of male fetuses and infants (McCarthy, 2019; Sanders and Stoecker, 2015; Kraemer, 2000), we find larger impacts on younger male than female siblings.

We also observe that the impact on respiratory hospitalizations appears to be monotonically decreasing with birth spacing—that is, younger siblings in families with a shorter birth spacing period experience larger differential impacts on hospitalizations in the first year of life. This pattern is consistent with the descriptive evidence presented in Figure 3, and speaks in favor of the mechanism of intra-family spread as being a key driver of respiratory disease among younger infant siblings. Further, these results suggest that our effects are *not* driven by differences in parental investments between older and younger siblings (and the potential interactions between these investments and our disease indices). As documented by Price (2008) in the U.S. setting, there are important differences in parent-child quality time between first- and second-born children, but this difference is larger when the birth spacing gap is longer. Thus, our pattern is the opposite of what would be predicted if differential parental time investment were the main channel.

Lastly, we find that the effects on respiratory hospitalizations among younger siblings are larger in sibling pairs with a short birth spacing in which the older child is in a childcare center than in pairs in which the older child is not. This result provides further support for our hypothesized mechanism of spread—that the older sibling gets exposed to respiratory

[19] Because a large share of children are enrolled in a childcare center from age 3 onward (see Appendix Figure A1), we do not study heterogeneity along this margin for siblings with a birth spacing gap that is longer than 2 years (given that the older children are 3 or older when their younger siblings are one year old). Additionally, the heterogeneity by childcare enrollment analysis sample is limited to sibling pairs born between September 1995 and September 2013, which is the period of time covered by our childcare enrollment data.

disease while in group childcare, and then "brings it home" to their more vulnerable younger brother or sister.

In Figure 5, we examine the effects of respiratory disease exposure in infancy on respiratory hospitalizations at later ages. Here, we plot the coefficients and 95% confidence intervals on our key interaction treatment variable from separate models that use as outcomes the annual number of respiratory disease hospitalizations, measured at different ages denoted on the $x-$axis. The graphs indicate that the large differential effect on hospitalizations before age one among younger siblings dissipates as they age. If anything, it appears that respiratory disease exposure in the first year of life is associated with a reduction in the number of overall respiratory hospitalizations at ages 3 to 4. These findings are consistent with the immunity formation hypothesis (Holt and Jones, 2000; M'Rabet et al., 2008; Côté et al., 2010; Fink et al., 2021), at least for some respiratory conditions. Moreover, our findings suggest that any differential effects on long-term educational and economic outcomes of early-life respiratory disease exposure among younger siblings are *not* driven by worse respiratory health in later childhood.

In Appendix Table A5, we explore the extent to which RSV contributes to the overall impact of respiratory disease. The table is identical to Table 2, except that we study the number of hospitalizations during the child's first year of life with an RSV diagnosis as the outcome (using cohorts born in 1994 or later), and we use an RSV-specific index instead of an index capturing all respiratory-related hospitalizations. We estimate that an additional RSV hospitalization per 100 children aged 13–71 months in a municipality increases a younger child's number of RSV hospitalizations in the first year of life by an average of 0.044 more than their older sibling's RSV hospitalizations at the same age. Moving from the 25th to the 75th percentile of the RSV index distribution amounts to a 0.005 differential increase in the number of RSV hospitalizations, or 27.8 percent at the sample mean. In Appendix Figure A4, we analyze the differential effects on RSV hospitalizations by age. Unlike what we saw for overall respiratory hospitalizations in Figure 5, we do not see a reduction in RSV hospitalizations at older ages, which makes sense as a prior RSV infection provides limited protection against future infections (Lambert et al., 2014; Fuentes et al., 2016). Thus, not all respiratory infections provide benefits in terms of immunity formation, and early-life RSV exposure may be an important driver of the adverse impacts on long-term outcomes that we describe next.

4.2 Long-Term Effects of Infancy Respiratory Disease Exposure on Educational and Economic Outcomes

Having established that local respiratory disease exposure among slightly older children predicts children's own hospitalizations for respiratory conditions before age one, and that this effect is much larger for younger relative to older siblings, we proceed to analyze children's long-term educational and economic outcomes.

We first present results using as outcomes the standardized 9th grade Danish and mathematics test scores in Tables 3 and 4, respectively. We find that an additional respiratory hospitalization in the municipality per 100 children aged 13–71 months reduces the 9th grade Danish test score by about 0.008 of a standard deviation more for younger siblings than older siblings, and this coefficient is marginally significant at the 10% level. The 25th to 75th percentile increase in the disease index amounts to an additional 0.013 of a standard deviation penalty on the Danish test score for the younger siblings relative to the older siblings. We do not observe a statistically significant impact on math test scores, although the interaction coefficient is similarly negative in sign.

Figure 6 presents results for the outcomes of high school and college graduation, by ages 18–32, respectively. We plot the coefficients and 95% confidence intervals on our key interaction term from separate models that use outcomes measured at the ages listed on the $x-$axis as dependent variables. For both high school and college graduation (Panels (a) and (b), respectively), consistent negative impacts are noticeable starting at the age when these outcomes can be affected (i.e., starting around age 19 for high school graduation, and age 23 for college graduation). For high school graduation, there appears to be some "catch-up" with age, suggesting that part of the overall effect stems from a delay in high school completion rather than a reduction in ever completing high school. The magnitude of the negative effect on college graduation, on the other hand, is quite stable between ages 24 and 32, although not always statistically significant at any given age.

Figure 7 reports results for our three main labor market outcomes: labor force participation, log income (conditional on being employed), and relative income rank, all measured at ages 18–32. As with the educational outcomes, we show the interaction term from separate regression models with outcomes measured at the ages listed on the $x-$axis. Panel (a) shows that the effects on labor force participation are not statistically significant at any age measured. That said, the pattern of coefficients suggests that there could be some positive impacts in early adulthood, followed by some negative impacts in late 20s, which may reflect the shift away from (or delay in) higher education that we just discussed.

When we analyze income among those who are employed in Panel (b) of Figure 7, we find an adverse differential effect of early life respiratory disease exposure among younger siblings, concentrated in their late 20s and early 30s. Pooling across income measured at ages 25–32, we estimate that an additional respiratory hospitalization per 100 children aged 13–71 months in an individual's municipality in the first year of life is associated with a 0.5 percent decline in income (see Table 5).[20] The 25th to 75th percentile effect size amounts to a 0.7 percent reduction in income at these ages. Similarly, we see a negative differential effect on relative income rank in Panel (c) of Figure 7. Here, the effect appears to materialize at younger ages and persist into the 30s.

Figure 8 explores the distributional impacts of early life exposure to respiratory illness further. We show coefficients and 95% confidence intervals on our key interaction treatment variable from models that use as outcomes indicators for being in different bins of the Danish income distribution within each birth cohort (where income is measured over ages 25–32): the 1–10th percentiles, the 11–25th percentiles, the 26–50th percentiles, the 51–75th percentiles, the 76–90th percentiles, and the 91–100th percentiles. We find a shift down from the top of the distribution: younger siblings exposed to more respiratory disease in the first year of life are significantly less likely to be in the top decile of the Danish income distribution. The coefficient for being in the 76–90th percentiles is also negative, but not statistically significant. At the same time, we see positive coefficients on the likelihoods of being in the lowest three bins of the income distribution.

The distributional impacts that we find differ somewhat from those identified in prior

[20] Table 5 reports results from our baseline models. Here, we use data at the person-by-age level, and study the outcome at ages 25–32. These models include age fixed effects and cluster standard errors on the municipality and individual level.

research on other types of early childhood shocks. For example, Isen et al. (2017b) find that reduced exposure to air pollution in the first year of life is associated with a shift from the bottom to the middle of the earnings distribution among US adults. While there are many mechanisms that could account for the difference in these patterns, one possibility is that the early life shock that we study—exposure to common respiratory viruses in infancy—is more universally prevalent across families with from different socio-economic backgrounds than a shock like air pollution, which disproportionately affects disadvantaged populations. Thus, our results suggest that even for children born in families that are relatively protected from adverse shocks due to their advantaged position in society, severe respiratory illness in early infancy can lower the likelihood that they end up at the top of the income distribution as adults. At the same time, as discussed more below, the magnitudes of our long-run effects are smaller than those documented in prior work, which may reflect both the lower severity of the shock that we study and the more advantaged population that it impacts.

Appendix Figures A5 and A6 explore heterogeneity in the long-run effects on educational and labor market outcomes. To reduce the number of estimates, we do not study outcomes at individual ages, and instead analyze high school and college graduation by age 30, and average labor market outcomes across ages 25–32. For the latter set of models, we use data at the person-by-age level, include age fixed effects, and cluster standard errors on the municipality and individual level. As with the heterogeneity analysis studying respiratory hospitalizations before age one in Figure 4, we use interaction models, in which subgroup indicators are interacted with the younger sibling indicator, the disease index, and the younger sibling indicator $\times$ disease index interaction. We then plot the coefficients and 95% confidence intervals from estimates of the triple interaction terms. While we mostly do not see much statistically significant heterogeneity—perhaps due to the smaller sample sizes used in studying long-run outcomes—we do find some suggestive evidence of differential effects by the younger sibling's gender. Specifically, the adverse negative effects on adult income are larger for males than females, which echoes our result of larger increases in respiratory hospitalizations by age 1 among younger brothers compared to younger sisters.

4.3 Effects on Mental Health Care Outcomes in Adolescence and Young Adulthood

Figure 9 presents results for the mental health care utilization outcomes, observed at ages 16–26.[21] We again plot the coefficients and 95% confidence intervals on our key interaction term from separate models that use outcomes measured at the ages listed on the $x-$axis as dependent variables.

The first two panels of Figure 9 focus on hospitalizations involving a mental health diagnosis. We find mostly elevated rates of mental health-related hospitalizations from age 16 through 26, both in terms of the extensive margin (Panel a) and in terms of the total number of hospitalizations (Panel b). Positive coefficients are observed consistently between age 16 and 26 (with the exception of age 25), with an average increase of 0.025 percentage point in the likelihood of having any mental health-related hospitalizations in a given age between 16 and 26 for each additional respiratory hospitalization per 100 children aged 13–71 months in an individual's municipality in the first year of life (see Appendix Tables A14 and A15, column (1)).[22] The 25th to 75th percentile effect size at these ages is a 9.6 percent increase in the likelihood of having any mental health-related hospitalizations and a 12.0 percent increase in the annual number of mental health-related hospitalizations, relative to each outcome's sample mean.

Panels (c) and (d) of Figure 9 present results for psychiatrist visits, which are somewhat less precise and appear a little later than the hospitalization effects (around ages 18 to 19). However, significantly positive effects are observed at many ages and the overall effect pattern is consistent with the increase in mental health-related hospitalizations shown in the first two panels. The 25th to 75th percentile effect size corresponds to a 5.0 percent increase in psychiatrist visits (see Appendix Table A16, column 1).

Appendix Figure A7 presents the heterogeneity analysis of mental health care outcomes averaged over ages 16–26. We use the same approach as we do for studying labor market

[21] The more limited age range of mental health care outcomes as compared to educational and labor market outcomes stems from the fact that we observed psychiatrist visits for a more limited set of years.

[22] Appendix Tables A14 and A15, column (1), report results from our baseline model. Here we use data at the person-by-age level, and study each outcome at ages 16–26. These models include age fixed effects and cluster standard errors on the municipality and individual level.

outcomes—we use data at the person-by-age level, include age fixed effects, and cluster standard errors on the municipality and individual level. We present triple interaction coefficients and 95% confidence intervals on the interaction term between each subgroup indicator, the younger sibling indicator, and the respiratory disease index. We do not observe strong evidence of heterogeneous effects. There is some suggestive indication that the increases in mental health care utilization are stronger among female than male younger siblings, which is consistent with other research suggesting that women are more likely than men to seek mental health care (Pattyn et al., 2015).

4.4 Magnitudes

How do our estimated long-run effects on economic and mental health outcomes compare to those in the prior literature? As noted above, we find that moving from the 25th to the 75th percentile of the respiratory disease index distribution is associated with an additional 0.7 percent reduction in adult income for second-born children. This effect size is slightly lower than the earnings impact of an 8 percent reduction in birth weight (Black et al., 2007) or a 7 percent increase in ambient air pollution in one's year of birth (Isen et al., 2017b). It also corresponds to about half of the effect of *in utero* exposure to the 1918 Spanish Influenza pandemic (Almond, 2006) and one-fifth of the effect of *in utero* exposure to a maternal influenza infection that requires hospitalization (Schwandt, 2018).[23]

It is additionally helpful to compare our estimates to those found in studies evaluating policies that reduce disease prevalence in the population. For example, Bhalotra and Venkataramani (2015) find that moving from the 75th to the 25th percentile in the pneumonia infection rate following the introduction of sulfa drugs leads to a 2.1 percent increase in adult income among exposed cohorts. Atwood (2022) and Chuard et al. (2022) find that the introduction of universal childhood measles vaccine lead to a 1.7 to 2.7 percent increase in adult family income among cohorts who benefited from the vaccine. Bütikofer and Salvanes (2020) document a 0.8 percent increase in adult income for cohorts who were in school during and after a tuberculosis control campaign in Norwegian municipalities that had above-median pre-campaign

[23]Note that our estimates represent intent-to-treat effects as not every child gets sick in response to exposure to a higher respiratory disease index.

tuberculosis levels.

We can also benchmark our estimates against the literature on birth order. In seminal work, Black et al. (2005) find an earnings disadvantage of 1.2 to 4.2 percent for second-born siblings compared to those who are first-born. Our birth order effect is within this range—we find a 1.6 percent income penalty for younger compared to older siblings in regressions that exclude the interaction term between the respiratory disease index and the younger sibling indicator (see Columns (1) and (3) of Table 5). However, when the interaction term is included, the main effect of birth order substantially decreases in magnitude and becomes statistically insignificant. This result suggests that an important part of the overall birth order effect on income could be explained by the second-born child's higher vulnerability to respiratory disease during infancy.

The effects on mental health that we estimate echo conclusions of other work documenting impacts of fetal and early childhood shocks on later mental health outcomes. For example, Almond and Mazumder (2011) find that exposure to Ramadan *in utero* leads to a near doubling of the incidence of mental and learning disabilities in adulthood in Uganda, and increases the rate of psychological disabilities in adulthood by 63 percent in Iraq. Persson and Rossin-Slater (2018) use data from Sweden, and find that experiencing the death of a close maternal relative while *in utero* is associated with a 25 percent increase in the likelihood of using ADHD medications around age 10, as well as 13 and 8 percent increases in the likelihoods of using drugs to treat depression and anxiety, respectively, around age 35. Adhvaryu et al. (2019) use a nationally-representative survey from Ghana, and show that a one standard deviation increase in the price of cocoa in one's year of birth—which improves the economic circumstances of Ghanaian families in cocoa-producing regions—reduces the likelihood of severe mental distress in adulthood by 3 percentage points, or about 50 percent at the mean prevalence rate. Our results on mental health are thus within the range of these estimates from studies based on a variety of contexts and types of shocks.

4.5 Additional Results

We examine the sensitivity of our results on short- and long-run outcomes across different specifications and different ways of measuring respiratory disease exposure in Appendix Tables

A6 through A17. For tractability, as with the heterogeneity analyses, we study high school and college graduation by age 30, average labor market outcomes across ages 25–32, and average mental health care utilization outcomes at ages 16–26. For the latter two sets of models, we use data at the person-by-age level, include age fixed effects, and cluster standard errors on the municipality and individual level. Then, we estimate different versions of model (1). Column (1) of each table presents the baseline model in which we include municipality, birth year, and birth month fixed effects and family background controls. Column (2) adds municipality-specific linear time trends to account for differential trends in outcomes across municipalities, while column (3) adds mother fixed effects that eliminate potential bias from unobserved genetic and family characteristics common among siblings.

In our baseline analysis, our respiratory disease index is based on the number of hospitalizations with a primary diagnosis of a respiratory condition. Columns (4) and (5) check the robustness of the results to alternative ways of constructing the disease index. Column (4) calculates the disease index based on number of hospitalizations including both primary and non-primary diagnoses for respiratory conditions, while in column (5) we construct it based on the number of children with at least one primary respiratory disease diagnosis (i.e., we count the number of children rather than the total number of hospitalizations). Our results on respiratory hospitalizations before age one are highly robust across these different modeling choices. Moreover, the effects on long-run educational, economic, and mental health outcomes in Appendix Tables A7 through A17 are largely consistent in terms of coefficient signs and magnitudes, although not every specification yields a statistically significant result for each outcome.

5 Conclusion

Respiratory illnesses are very common among young children, especially in families with more than one child. Despite their regular occurrence, there is limited population-level evidence on the role of intra-family transmission, or on the long-term causal impacts of exposure to endemic respiratory disease during infancy. This paper uses linked administrative data from Denmark spanning four decades to document the importance of birth order in driving susceptibility to respiratory infection. We find that younger siblings are two to three times more likely to be hospitalized for respiratory conditions during their first year of life compared to the older siblings at the same age, and this disparity is especially large when hospitalizations are measured in the first three months of life. Additional analyses of the seasonality in hospitalizations and heterogeneity across siblings with different birth spacing gaps point to the importance of intra-family transmission in explaining this birth order effect: older children "bring home" common respiratory viruses (such as RSV), making their younger siblings susceptible to severe illness early in life.

We then combine the birth order variation with variation in local respiratory disease prevalence to study long-term effects of early-life disease on health, human capital, and economic outcomes. We show that exposure to severe respiratory illness during infancy has negative consequences on both educational and economic outcomes in adulthood. Our results show that moving from the 25th to the 75th percentile in the local respiratory disease prevalence distribution reduces the likelihood of on-time high school and college graduation, and leads to a 0.7 percent additional reduction in age 25–32 earnings for younger compared to older siblings. While we do not find that infancy exposure to respiratory disease adversely affects respiratory health at older ages, our analysis of mental health care outcomes suggests that impaired brain development could be an important channel driving the effects on human capital and labor market productivity. We find evidence of elevated rates of mental health-related hospitalizations and visits to psychiatrists at ages 16–26 resulting from infancy disease exposure among younger siblings.

The long-term effects that we estimate represent the overall net impacts of respiratory disease exposure during infancy. Thus, these estimates incorporate any potential benefits associated with increased immunity, as well as parental responses to the health shocks. In sum, our findings suggest that policies mitigating the spread of respiratory diseases among young children may have large long-term benefits, which are likely not incorporated into current cost-benefit evaluations.

References

Adams-Chapman, Ira and Barbara J Stoll, "Neonatal infection and long-term neurodevelopmental outcome in the preterm infant," *Current opinion in infectious diseases*, 2006, *19* (3), 290–297.

Adhvaryu, Achyuta, James Fenske, and Anant Nyshadham, "Early life circumstance and adult mental health," *Journal of Political Economy*, 2019, *127* (4), 1516–1549.

Almond, Douglas, "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population," *Journal of Political Economy*, 2006, *114* (4), 672–712.

_ **and Bhashkar Mazumder**, "The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data," *American Economic Review*, 2005, *95* (2), 258–262.

_ **and** _ , "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy," *American Economic Journal: Applied Economics*, 2011, *3* (4), 56–85.

_ **, Janet Currie, and Valentina Duque**, "Childhood circumstances and adult outcomes: Act II," *Journal of Economic Literature*, 2018, *56* (4), 1360–1446.

_ **, Lena Edlund, and Mårten Palme**, "Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden," *The Quarterly journal of economics*, 2009, *124* (4), 1729–1772.

Atwood, Alicia, "The long-term effects of measles vaccination on earnings and Employment," *American Economic Journal: Economic Policy*, 2022, *14* (2), 34–60.

Bailey, Martha J, Hilary W Hoynes, Maya Rossin-Slater, and Reed Walker, "Is the social safety net a long-term investment? Large-scale evidence from the food stamps program," Working Paper w26942, National Bureau of Economic Research 2020.

Barker, David J, "The fetal and infant origins of adult disease.," *BMJ: British Medical Journal*, 1990, *301* (6761), 1111.

Barofsky, Jeremy, Tobenna D Anekwe, and Claire Chase, "Malaria eradication and economic outcomes in sub-Saharan Africa: evidence from Uganda," *Journal of health economics*, 2015, *44*, 118–136.

Barreca, Alan I, "The long-term economic impact of in utero and postnatal exposure to malaria," *Journal of Human resources*, 2010, *45* (4), 865–892.

Bateson, Patrick, Peter Gluckman, and Mark Hanson, "The biology of developmental plasticity and the Predictive Adaptive Response hypothesis," *The Journal of physiology*, 2014, *592* (11), 2357–2368.

Beuchert, Louise Voldby, Maria Knoth Humlum, and Rune Vejlin, "The length of maternity leave and family health," *Labour Economics*, 2016, *43*, 55–71.

Bhalotra, Sonia R and Atheendar Venkataramani, "Cognitive development and infectious disease: Gender differences in investments and outcomes," *Available at SSRN 2372542*, 2013.

— **and** — , "Shadows of the captain of the men of death: Early life health interventions, human capital investments, and institutions," *Human Capital Investments, and Institutions (August 8, 2015)*, 2015.

Biasi, Barbara, Michael S Dahl, and Petra Moser, "Career effects of mental health," Working Paper w29031, National Bureau of Economic Research 2021.

Bilbo, Staci D and Jaclyn M Schwarz, "The immune system and developmental programming of brain and behavior," *Frontiers in neuroendocrinology*, 2012, *33* (3), 267–286.

Black, Maureen M, Susan P Walker, Lia CH Fernald, Christopher T Andersen, Ann M DiGirolamo, Chunling Lu, Dana C McCoy, Günther Fink, Yusra R Shawar, and Jeremy Shiffman, "Early childhood development coming of age: science through the life course," *The Lancet*, 2017, *389* (10064), 77–90.

Black, Sandra E, Aline Bütikofer, Paul J Devereux, and Kjell G Salvanes, "This is only a test? Long-run and intergenerational impacts of prenatal exposure to radioactive fallout," *Review of Economics and Statistics*, 2019, *101* (3), 531–546.

— **, Paul J Devereux, and Kjell G Salvanes**, "The more the merrier? The effect of family size and birth order on children's education," *The Quarterly Journal of Economics*, 2005, *120* (2), 669–700.

— **,** — **, and** — , "From the cradle to the labor market? The effect of birth weight on adult outcomes," *The Quarterly Journal of Economics*, 2007, *122* (1), 409–439.

— **,** — **, and** — , "Does grief transfer across generations? Bereavements during pregnancy and child outcomes," *American Economic Journal: Applied Economics*, 2016, *8* (1), 193–223.

— **, Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen**, "Sibling spillovers," *The Economic Journal*, 2021, *131* (633), 101–128.

Bleakley, Hoyt, "Malaria eradication in the Americas: A retrospective analysis of childhood exposure," *American Economic Journal: Applied Economics*, 2010, *2* (2), 1–45.

Breining, Sanni, Joseph Doyle, David N Figlio, Krzysztof Karbownik, and Jeffrey Roth, "Birth order and delinquency: Evidence from Denmark and Florida," *Journal of Labor Economics*, 2020, *38* (1), 95–142.

Brenøe, Anne and Ramona Molitor, "Birth Order and Health of Newborns: What Can We Learn from Danish Registry Data?," *Journal of Population Economics*, 2018, *31*, 363–395.

Buckles, Kasey and Shawna Kolka, "Prenatal investments, breastfeeding, and birth order," *Social Science & Medicine*, 2014, *118*, 66–70.

Bütikofer, Aline and Kjell G Salvanes, "Disease control and inequality reduction: Evidence from a tuberculosis testing and vaccination campaign," *The Review of Economic Studies*, 2020, *87* (5), 2087–2125.

— , **Christopher J Cronin, and Meghan M Skira**, "Employment effects of healthcare policy: Evidence from the 2007 FDA black box warning on antidepressants," *Journal of Health Economics*, 2020, *73*, 102348.

Carbonell-Estrany, Xavier, Eduardo G Pérez-Yarza, Laura Sanchez García, Juana M Guzmán Cabañas, Elena Villarrubia Bòria, Belén Bernardo Atienza, and IRIS (Infección Respiratoria Infantil por Virus Respiratorio Sincitial) Study Group, "Long-term burden and respiratory effects of respiratory syncytial virus hospitalization in preterm infants—the SPRING study," *PloS one*, 2015, *10* (5), e0125422.

Chang, Simon, Belton Fleisher, Seonghoon Kim, and Shi yung Liu, "Long-Term Health Effects of Malaria Exposure around Birth: Evidence from Colonial Taiwan," *Economic Development and Cultural Change*, 2014, *62* (3), 519–536.

Chuard, Caroline, Hannes Schwandt, Alexander D Becker, and Masahiko Haraguchi, "Economic vs. Epidemiological Approaches to Measuring the Human Capital Impacts of Infectious Disease Elimination," Technical Report, National Bureau of Economic Research 2022.

Côté, Sylvana M, Amélie Petitclerc, Marie-France Raynault, Qian Xu, Bruno Falissard, Michel Boivin, and Richard E Tremblay, "Short-and long-term risk of infections as a function of group child care attendance: an 8-year population-based study," *Archives of pediatrics & adolescent medicine*, 2010, *164* (12), 1132–1137.

Cowling, Benjamin J, Sheikh Taslim Ali, Tiffany WY Ng, Tim K Tsang, Julian CM Li, Min Whui Fong, Qiuyan Liao, Mike YW Kwan, So Lun Lee, Susan S Chiu et al., "Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study," *The Lancet Public Health*, 2020, *5* (5), e279–e288.

Currie, Janet and Douglas Almond, "Human capital development before age five," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1315–1486.

Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl, "Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India," *American Economic Journal: Applied Economics*, 2010, *2* (2), 72–94.

Daysal, N Meltem, Marianne Simonsen, Mircea Trandafir, and Sanni Breining, "Spillover effects of early-life medical interventions," *Review of Economics and Statistics*, 2020, pp. 1–46.

den Berg, Gerard J Van and Bettina M Siflinger, "The effects of day care on health during childhood: evidence by age," Discussion Paper DP15036, CEPR 2020.

Deverman, Benjamin E and Paul H Patterson, "Exogenous leukemia inhibitory factor stimulates oligodendrocyte progenitor cell proliferation and enhances hippocampal remyelination," *Journal of Neuroscience*, 2012, *32* (6), 2100–2109.

Eppig, Christopher, Corey L Fincher, and Randy Thornhill, "Parasite prevalence and the worldwide distribution of cognitive ability," *Proceedings of the Royal Society B: Biological Sciences*, 2010, *277* (1701), 3801–3808.

Fink, Günther, Atheendar S Venkataramani, and Arianna Zanolini, "Early life adversity, biological adaptation, and human capital: evidence from an interrupted malaria control program in Zambia," *Journal of Health Economics*, 2021, *80*, 102532.

Fuentes, Sandra, Elizabeth M Coyle, Judy Beeler, Hana Golding, and Surender Khurana, "Antigenic fingerprinting following primary RSV infection in young children identifies novel antigenic sites and reveals unlinked evolution of human antibody repertoires to fusion and attachment glycoproteins," *PLoS pathogens*, 2016, *12* (4), e1005554.

Gensowski, Miriam, Torben Heien Nielsen, Nete Munk Nielsen, Maya Rossin-Slater, and Miriam Wüst, "Childhood health shocks, comparative advantage, and long-term outcomes: Evidence from the last Danish polio epidemic," *Journal of health economics*, 2019, *66*, 27–36.

Gluckman, Peter and Mark Hanson, *Mismatch: Why our world no longer fits our bodies*, Oxford University Press, 2006.

Haan, Monique De, "Birth order, family size and educational attainment," *Economics of Education Review*, 2010, *29* (4), 576–588.

Holt, PG and CA Jones, "The development of the immune system during pregnancy and early life," *Allergy*, 2000, *55* (8), 688–697.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond, "Long-run impacts of childhood access to the safety net," *American Economic Review*, 2016, *106* (4), 903–34.

Isen, Adam, Maya Rossin-Slater, and Reed Walker, "Relationship between season of birth, temperature exposure, and later life wellbeing," *Proceedings of the National Academy of Sciences*, 2017, *114* (51), 13447–13452.

_ , _ , and W Reed Walker, "Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970," *Journal of Political Economy*, 2017, *125* (3), 848–902.

Kelly, Elaine, "The scourge of asian flu in utero exposure to pandemic influenza and the development of a cohort of british children," *Journal of Human resources*, 2011, *46* (4), 669–694.

Kneyber, MCJ, EW Steyerberg, R De Groot, and HA Moll, "Long-term effects of respiratory syncytial virus (RSV) bronchiolitis in infants and young children: a quantitative review," *Acta Paediatrica*, 2000, *89* (6), 654–660.

Korppi, M, E Piippo-Savolainen, K Korhonen, and S Remes, "Respiratory morbidity 20 years after RSV infection in infancy," *Pediatric pulmonology*, 2004, *38* (2), 155–160.

Kraemer, Sebastian, "The fragile male," *Bmj*, 2000, *321* (7276), 1609–1612.

Kuecken, Maria, Josselin Thuilliez, and Marie-Anne Valfort, "Disease and Human Capital Accumulation: Evidence from the Roll Back Malaria Partnership in Africa," *The Economic Journal*, 2021, *131* (637), 2171–2202.

Kusel, Merci MH, Nicholas H de Klerk, Tatiana Kebadze, Vaike Vohma, Patrick G Holt, Sebastian L Johnston, and Peter D Sly, "Early-life respiratory viral infections, atopic sensitization, and risk of subsequent development of persistent asthma," *Journal of Allergy and Clinical Immunology*, 2007, *119* (5), 1105–1110.

Lambert, Laura, Agnes M Sagfors, Peter JM Openshaw, and Fiona J Culley, "Immunity to RSV in early-life," *Frontiers in immunology*, 2014, *5*, 466.

Lehmann, Jee-Yeon K, Ana Nuevo-Chiquero, and Marian Vidal-Fernandez, "The early origins of birth order differences in children's outcomes and parental behavior," *Journal of Human Resources*, 2018, *53* (1), 123–156.

Leung, Nancy HL, Daniel KW Chu, Eunice YC Shiu, Kwok-Hung Chan, James J McDevitt, Benien JP Hau, Hui-Ling Yen, Yuguo Li, Dennis KM Ip, JS Malik Peiris et al., "Respiratory virus shedding in exhaled breath and efficacy of face masks," *Nature medicine*, 2020, *26* (5), 676–680.

Lin, Ming-Jen and Elaine M Liu, "Does in utero exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment," *Journal of health economics*, 2014, *37*, 152–163.

Lucas, Adrienne M, "Malaria eradication and educational attainment: evidence from Paraguay and Sri Lanka," *American Economic Journal: Applied Economics*, 2010, *2* (2), 46–71.

Lynge, Elsebeth, Jakob Lynge Sandegaard, and Matejka Rebolj, "The Danish national patient register," *Scandinavian journal of public health*, 2011, *39* (7_suppl), 30–33.

McCarthy, Margaret M, "Stress during pregnancy: Fetal males pay the price," *Proceedings of the National Academy of Sciences*, 2019, *116* (48), 23877–23879.

M'Rabet, Laura, Arjen Paul Vos, Gunther Boehm, and Johan Garssen, "Breastfeeding and its role in early development of the immune system in infants: consequences for health later in life," *The Journal of nutrition*, 2008, *138* (9), 1782S–1790S.

O'Shea, T Michael, Bhavesh Shah, Elizabeth N Allred, Raina N Fichorova, Karl CK Kuban, Olaf Dammann, Alan Leviton, ELGAN Study Investigators et al., "Inflammation-initiating illnesses, inflammation-related proteins, and cognitive impairment in extremely preterm infants," *Brain, behavior, and immunity*, 2013, *29*, 104–112.

Pattyn, Elise, Mieke Verhaeghe, and Piet Bracke, "The gender gap in mental health service use," *Social psychiatry and psychiatric epidemiology*, 2015, *50*, 1089–1095.

Pedersen, Carsten Bøcker, "The Danish civil registration system," *Scandinavian journal of public health*, 2011, *39* (7_suppl), 22–25.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt, "Poorly measured confounders are more useful on the left than on the right," *Journal of Business & Economic Statistics*, 2019, *37* (2), 205–216.

Persson, Petra and Maya Rossin-Slater, "Family ruptures, stress, and the mental health of the next generation," *American economic review*, 2018, *108* (4-5), 1214–52.

Petersson, Flemming, Mikkel Baadsgaard, and Lau Caspar Thygesen, "Danish registers on personal labour market affiliation," *Scandinavian journal of public health*, 2011, *39* (7_suppl), 95–98.

Pichler, Stefan and Nicolas R Ziebarth, "Labor market effects of US sick pay mandates," *Journal of Human Resources*, 2020, *55* (2), 611–659.

Pitzer, Virginia E, Cécile Viboud, Wladimir J Alonso, Tanya Wilcox, C Jessica Metcalf, Claudia A Steiner, Amber K Haynes, and Bryan T Grenfell, "Environmental drivers of the spatiotemporal dynamics of respiratory syncytial virus in the United States," *PLoS pathogens*, 2015, *11* (1), e1004591.

Price, Joseph, "Parent-child quality time does birth order matter?," *Journal of human resources*, 2008, *43* (1), 240–265.

Pruckner, Gerald J., Nicole Schneeweis, Thomas Schober, and Martina Zweimüller, "Birth order, parental health investment, and health in childhood," *Journal of Health Economics*, 2021, *76*, 102426.

Rasmussen, Astrid Würtz, "Increasing the length of parents' birth-related leave: The effect on children's long-term educational outcomes," *Labour Economics*, 2010, *17* (1), 91–100.

Régnier, Stéphane A and Jasper Huels, "Association between respiratory syncytial virus hospitalizations in infants and respiratory sequelae: systematic review and meta-analysis," *The Pediatric infectious disease journal*, 2013, *32* (8), 820–826.

Rha, Brian, Aaron T Curns, Joana Y Lively, Angela P Campbell, Janet A Englund, Julie A Boom, Parvin H Azimi, Geoffrey A Weinberg, Mary A Staat, Rangaraj Selvarangan et al., "Respiratory syncytial virus–associated hospitalizations among young children: 2015–2016," *Pediatrics*, 2020, *146* (1).

Sanders, Nicholas J and Charles Stoecker, "Where have all the young men gone? Using sex ratios to measure fetal death rates," *Journal of health economics*, 2015, *41*, 30–45.

Schwandt, Hannes, "The lasting legacy of seasonal influenza: In-utero exposure and labor market outcomes," *CEPR Discussion Paper*, 2018.

van den Berg, Gerard J, Stephanie von Hinke, and Nicolai Vitt, "Early life exposure to measles and later-life outcomes: Evidence from the introduction of a vaccine," Technical Report, School of Economics, University of Bristol, UK 2023.

Venkataramani, Atheendar S, "Early life exposure to malaria and cognition in adulthood: evidence from Mexico," *Journal of health economics*, 2012, *31* (5), 767–780.

Viinikainen, Jutta, Alex Bryson, Petri Böckerman, Marko Elovainio, Nina Hutri-Kähönen, Markus Juonala, Terho Lehtimäki, Katja Pahkala, Suvi Rovio, Laura Pulkki-Råback et al., "Do childhood infections affect labour market outcomes in adulthood and, if so, how?," *Economics & Human Biology*, 2020, *37*, 100857.

Vliegenthart, Roseanne JS, Wes Onland, Aleid G van Wassenaer-Leemhuis, Anne PM De Jaegere, Cornelieke SH Aarnoudse-Moens, and Anton H van Kaam, "Restricted ventilation associated with reduced neurodevelopmental impairment in preterm infants," *Neonatology*, 2017, *112* (2), 172–179.

White, Corey, "Measuring social and externality benefits of influenza vaccination," *Journal of Human Resources*, 2019, pp. 1118–9893R2.

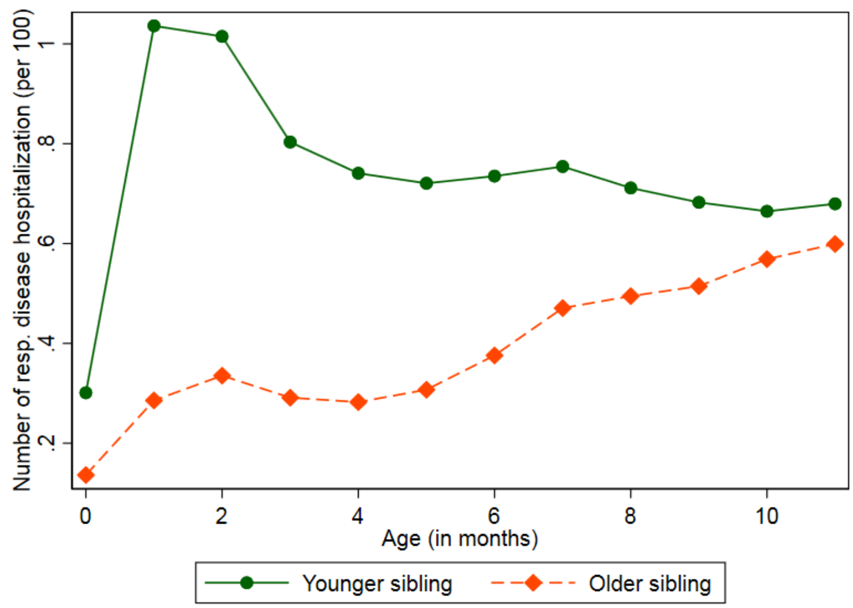
Yi, Junjian, James J Heckman, Junsen Zhang, and Gabriella Conti, "Early health shocks, intra-household resource allocation and child outcomes," *The Economic Journal*, 2015, *125* (588), F347–F371.

Zomer-Kooijker, Kim, Cornelis K van der Ent, Marieke JJ Ermers, Cuno SPM Uiterwaal, Maroeska M Rovers, Louis J Bont, and RSV Corticosteroid Study Group, "Increased risk of wheeze and decreased lung function after respiratory syncytial virus infection," *PloS one*, 2014, *9* (1), e87162.

6 Figures

Figure 1: Number of Respiratory Hospitalizations per 100 Children, by Child Age in Months, Older versus Younger Siblings

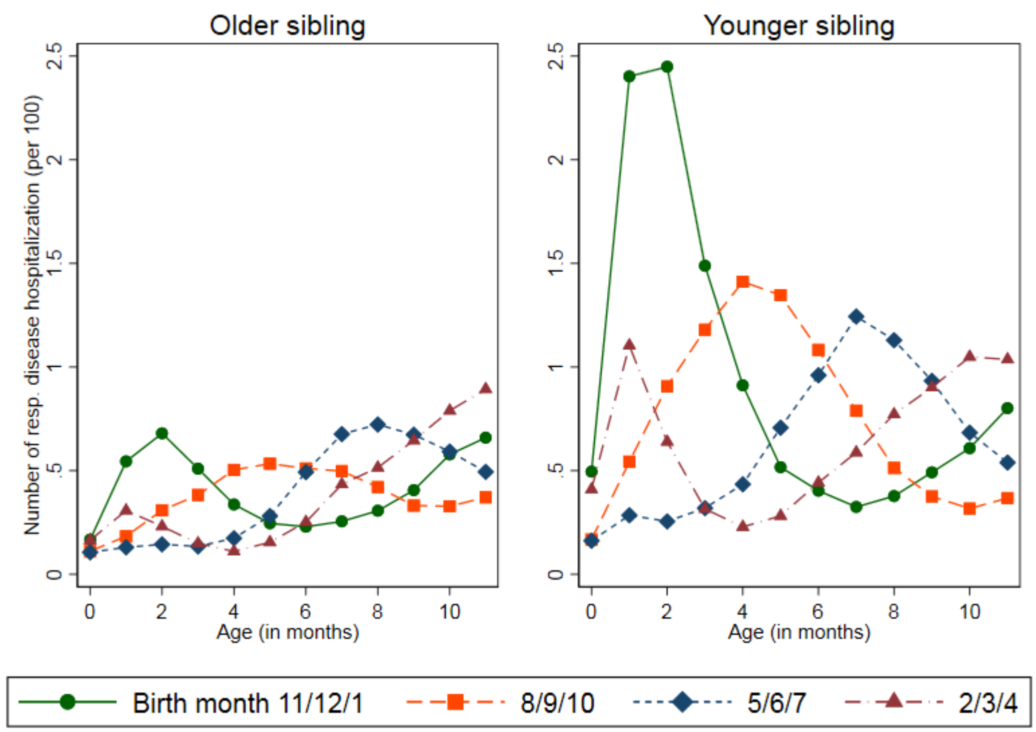
(a) During First Year of Life



(b) During First Five Years of Life

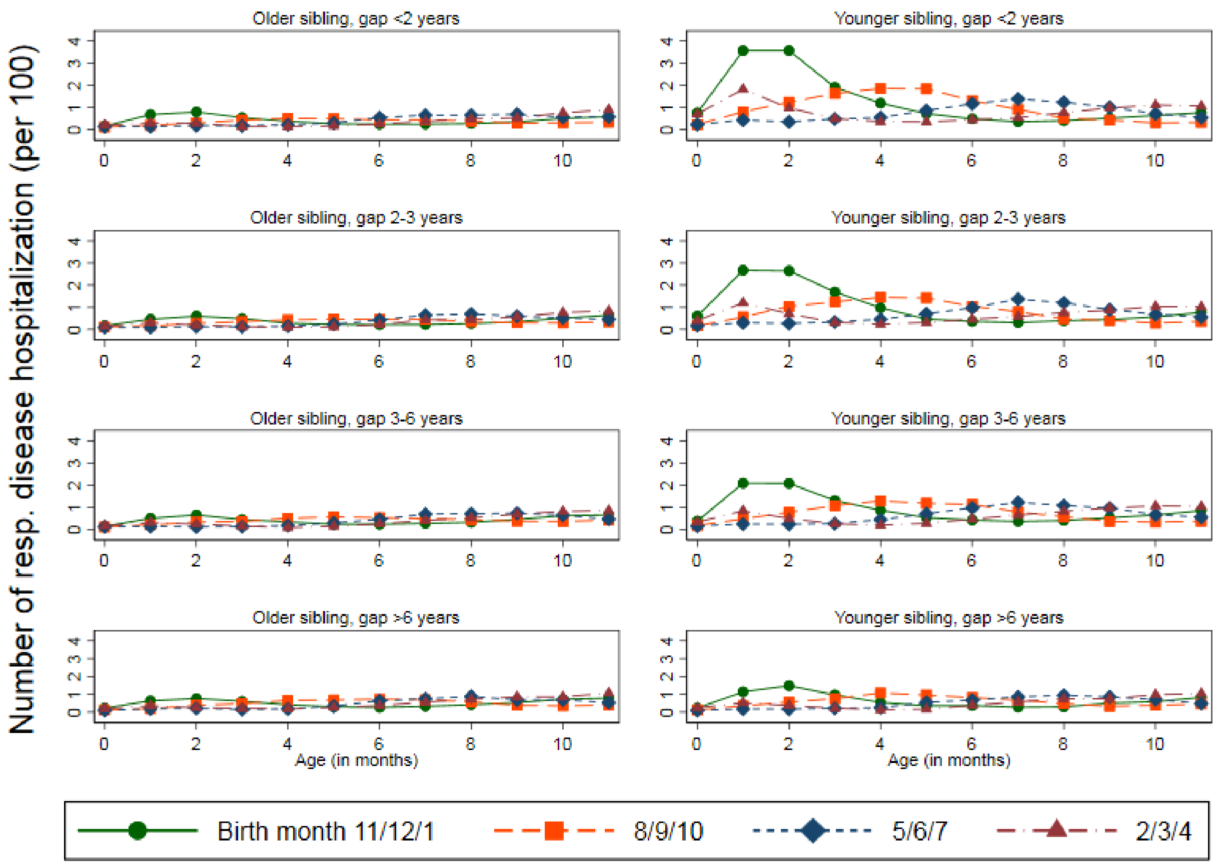
Notes: These figures plot the number of hospitalizations with respiratory illness diagnoses (per 100 children) by month of age, separately for older and younger siblings in our data.

Figure 2: Number of Respiratory Hospitalizations per 100 Children, by Child Age in Months and Season of Birth, Older versus Younger Siblings



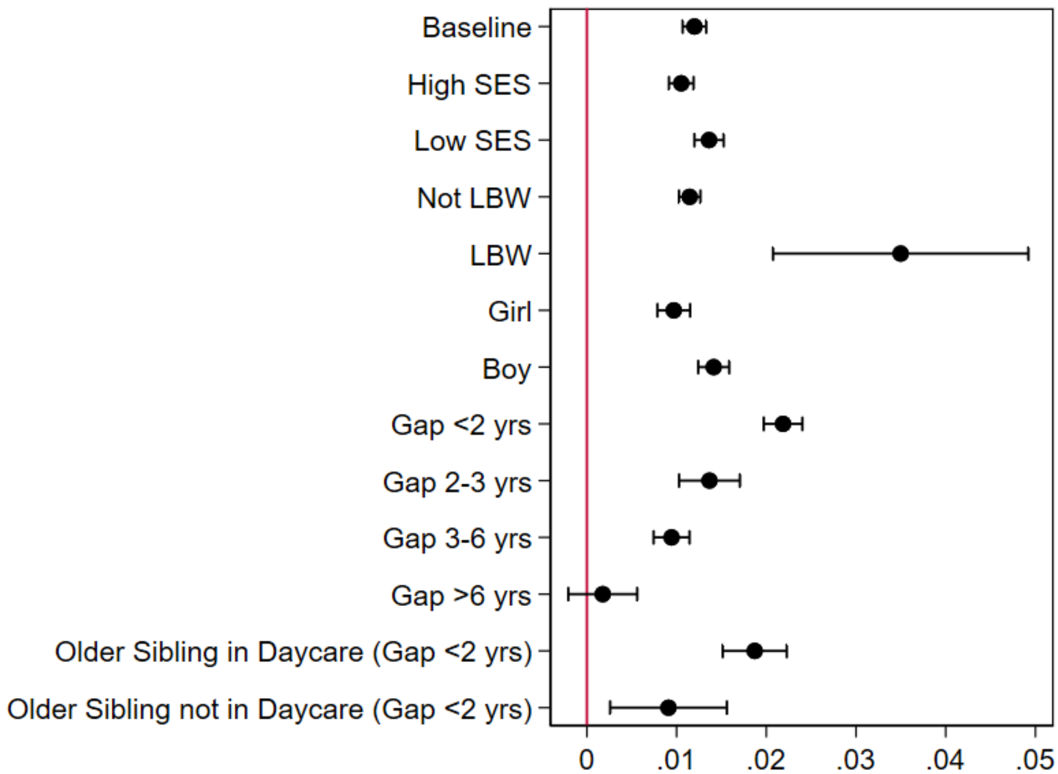
Notes: These figures plot the number of hospitalizations with respiratory illness diagnoses (per 100 children) by month of age and by the season of birth of the child, separately for older and younger siblings in our data.

Figure 3: Number of Respiratory Hospitalizations per 100 Children, by Child Age in Months, Season of Birth, and Birth Spacing, Older versus Younger Siblings



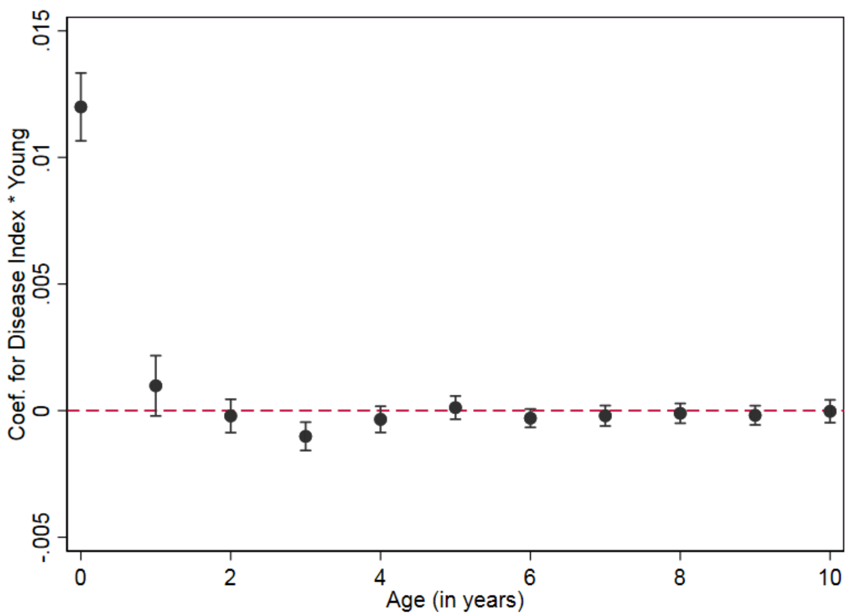
Notes: These figures plot the number of hospitalizations with respiratory illness diagnoses (per 100 children) by month of age and by the season of birth of the child, separately for older and younger siblings with different birth spacing gaps in our data.

Figure 4: Heterogeneous Effects of the Respiratory Disease Exposure Index on the Annual Number of Younger Siblings' Respiratory Hospitalizations



Notes: This figure plots the heterogeneity effects of the respiratory disease exposure on younger siblings respiratory hospitalizations among different sub-populations. The baseline coefficient and 95% confidence intervals are from the interaction term between the overall respiratory disease index and the younger sibling indicator from model (1). The respiratory disease exposure index is the number of inpatient admissions with any respiratory disease primary diagnosis among children aged 13–71 months per 100 children in each child's municipality of birth during the first year of life, excluding any hospitalizations of an older sibling. Effects by sub-groups are from 5 separate regressions: 1) high vs. low socioeconomic status (SES), grouped based on the mother's education level in the year of birth being above or below the median level among mothers in the same year; 2) low birth weight (LBW) status; 3) child gender; 4) birth spacing; and 5) whether the older child is in a childcare center during the first year of life of the younger child, restricting to sibling pairs born within 2 years of each other, and between September 1995 and September 2013 (the period of time covered by our childcare enrollment data). In each regression, the full set of sub-group indicators are interacted with the younger sibling indicator, the disease index, and the younger sibling indicator × disease index interaction. Coefficients and 95% confidence intervals of the triple interaction term are plotted accordingly. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls, including indicator for child gender, the sibling pair's birth spacing (in months) and the birth spacing interacted with the indicator for the younger child, mother's age and age squared, indicator for the mother being foreign-born, indicators for mother's education level (high school degree, college degree or higher), and an indicator for the parents being married or cohabiting at the time of childbirth. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

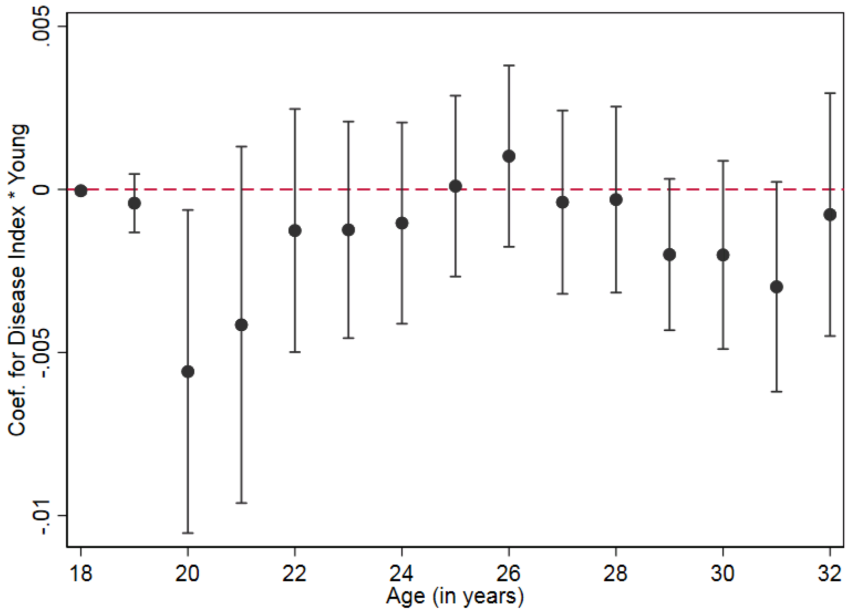
Figure 5: Effects of the Respiratory Disease Exposure Index on the Annual Number of Younger Siblings' Respiratory Hospitalizations, by Age of Observation



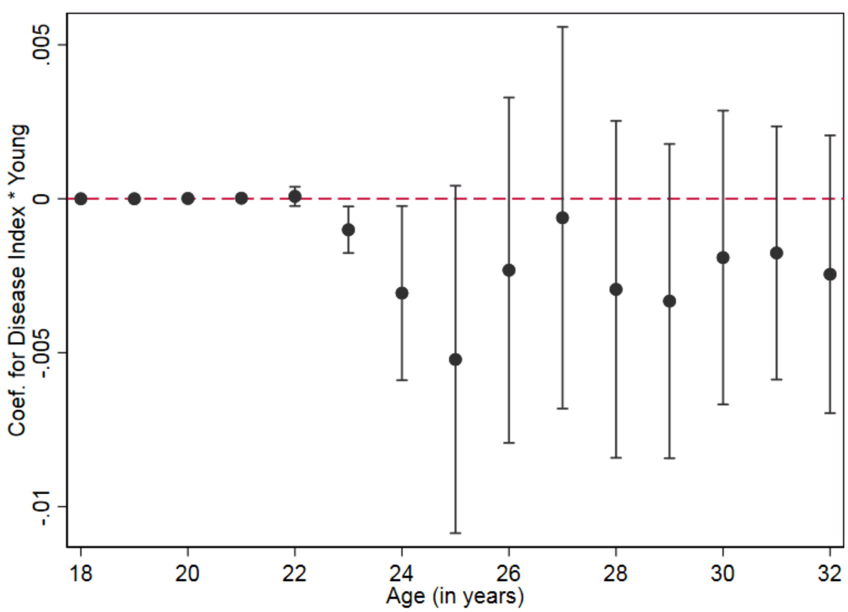
Notes: This figure plots the coefficients and 95% confidence intervals on the interaction term between the overall respiratory disease index and the younger sibling indicator from model (1), using as the outcome the annual number of hospitalizations with all respiratory diagnoses, measured at ages specified on the x-axis. The respiratory disease exposure index is the number of inpatient admissions with any respiratory disease primary diagnosis among children aged 13–71 months per 100 children in each child's municipality of birth during the first year of life, excluding any hospitalizations of an older sibling. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls, including indicator for child gender, the sibling pair's birth spacing (in months) and the birth spacing interacted with the indicator for the younger child, mother's age and age squared, indicator for the mother being foreign-born, indicators for mother's education level (high school degree, college degree or higher), and an indicator for the parents being married or cohabiting at the time of childbirth. . Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

Figure 6: Effects of the Respiratory Disease Exposure Index on Younger Siblings' Educational Outcomes, by Age of Observation

(a) High School Graduation



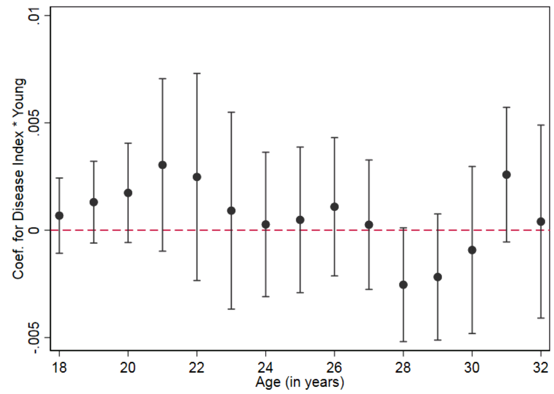
(b) College Graduation



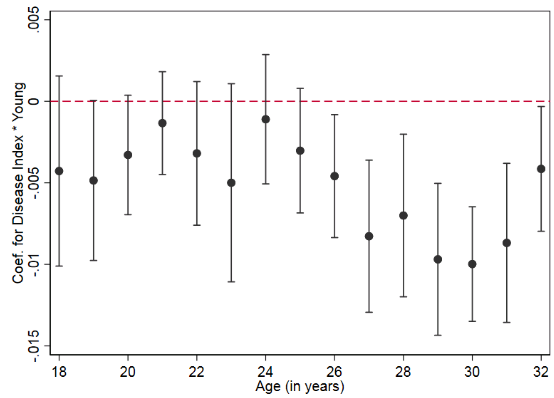
Notes: These figures plot the coefficients and 95% confidence intervals on the interaction term between the disease index and the younger sibling indicator from model (1), using outcomes measured at ages specified on the x-axes. At each age, we require both of the siblings are observed in the data. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 5 for more details about the specifications and variables. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

Figure 7: Effects of the Respiratory Disease Exposure Index on Younger Siblings' Labor Market Outcomes, by Age of Observation

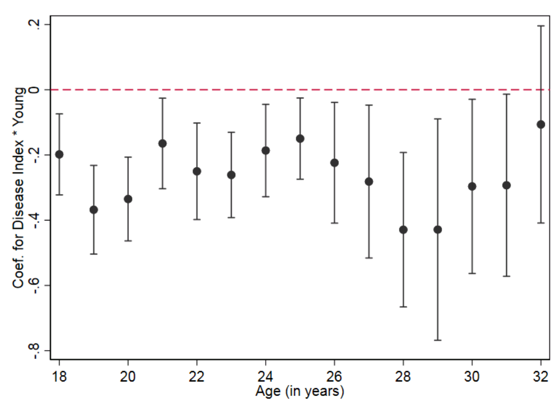
(a) Labor Force Participation



(b) Income (conditional on employed)

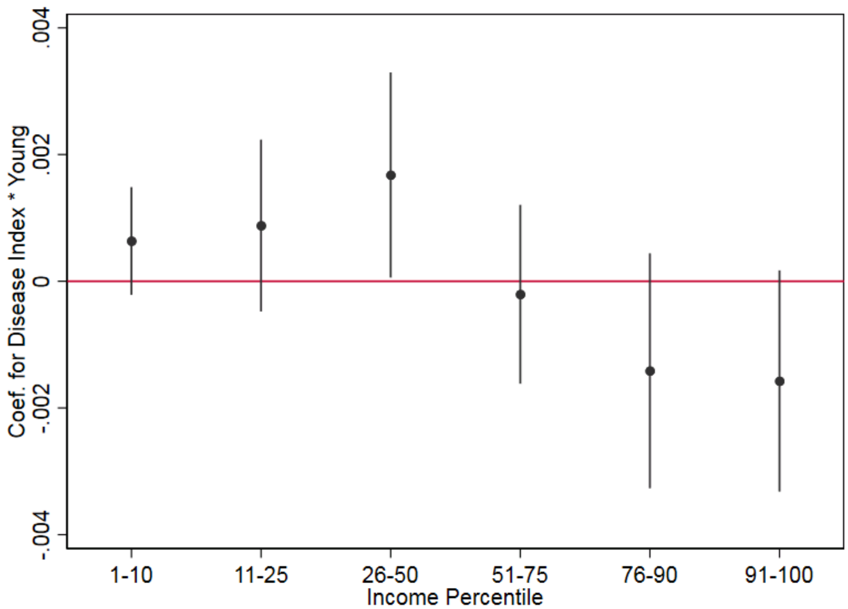


(c) Income Percentile



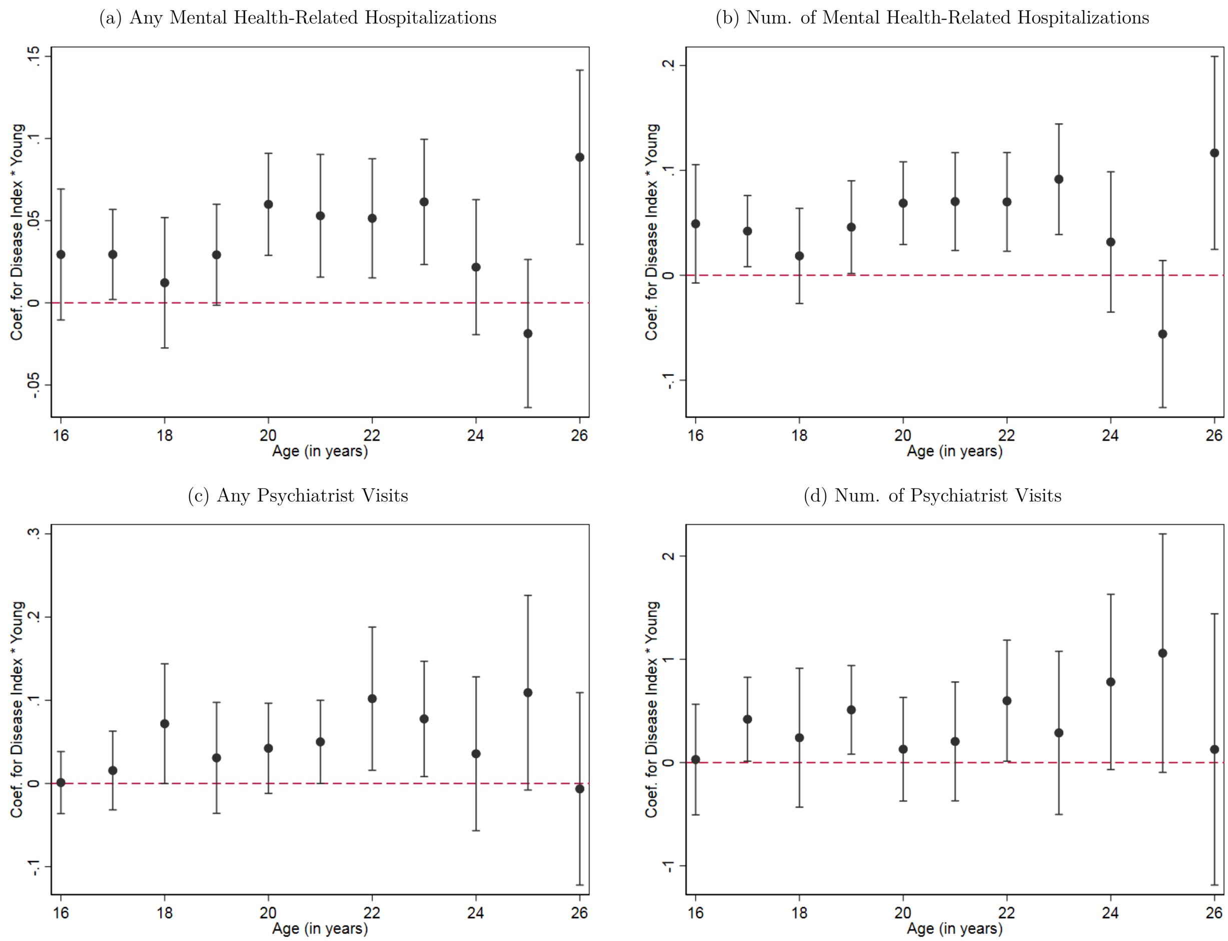
Notes: These figures plot the coefficients and 95% confidence intervals on the interaction term between the disease index and the younger sibling indicator from model (1), using outcomes measured at ages specified on the x-axes. At each age, we require both of the siblings are observed in the data. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 5 for more details about the specifications and variables. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

Figure 8: Effects of the Respiratory Disease Exposure Index on Younger Siblings' Income Distribution



Notes: This figure plots the coefficients and 95% confidence intervals on the interaction term between the disease index and the younger sibling indicator from model (1) with age fixed effects. The sample includes sibling pairs at age 25-32, with each observation at person-by-age level. The outcome is an indicator for the income percentile falling into each percentile bin denoted on the x-axis among population of the same age in the same year. All regressions include municipality, year of birth, month of birth, age fixed effects, and family background controls. Confidence intervals are constructed from two-way clustered standard errors at the individual and municipality of birth levels.

Figure 9: Effects of the Respiratory Disease Exposure Index on Younger Siblings' Mental Health Care Outcomes, by Age of Observation



Notes: These figures plot the coefficients and 95% confidence intervals on the interaction term between the disease index and the younger sibling indicator from model (1), using the mental health care outcomes measured at ages specified on the x-axes. At each age, we require both of the siblings are observed in the data. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 5 for more details about the specifications and variables. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

7 Tables

Table 1: Variable Means

	Older Siblings	Younger Siblings
Disease Exposure Indices (Per 100 Children Aged 13–71 Months)		
Respiratory Disease Exposure Index	2.792	2.876
Respiratory Disease Exposure Index (post-1993 cohorts)	3.026	3.016
RSV Exposure Index (post-1993 cohorts)	0.107	0.102
Child Characteristics		
Male Child	0.514	0.514
Birth Weight (grams)	3431.059	3588.840
Birth Spacing (months)	41.961	41.961
*Respiratory Disease Hospitalizations by Age 1 (*100)*		
Number of Respiratory Disease Hospitalizations by Age 1	4.638	8.955
Number of Respiratory Disease Hospitalizations by Age 1 (post-1993 cohorts)	4.457	9.888
Number of RSV Hospitalizations by Age 1 (post-1993 cohorts)	0.848	2.733
*Hospitalizations and Mental Health Outcomes, Ages 15-25 (*100)*		
Number of Mental Health-related Hospitalizations	0.412	0.486
Number of Psychiatrist Visits	6.381	6.451
Any Mental Health-related Hospitalizations	0.353	0.403
Any Psychiatrist Visit	0.893	0.933
Educational and Labor Market Outcomes		
High School Degree, Age 30	0.849	0.840
College Degree, Age 30	0.447	0.427
Danish Test Score, Grade 9	0.152	0.048
Math Test Score, Grade 9	0.207	0.077
Log Income (conditional on employed), Age 25-32	10.901	10.871
Income Percentile, Age 25-32	56.925	55.119
In Labor Force, Age 25-32	0.650	0.647
Family Background Characteristics		
Mother's Age at Childbirth	26.819	30.318
Mother is Foreign-Born	0.045	0.045
Mother has High School Degree	0.750	0.789
Mother has College Degree	0.302	0.367
Parents are Married/Cohabiting (Year after birth)	0.937	0.950
Log Household Income	11.424	11.600
Observations	581991	581991

Notes: This table presents the means of key variables in our analysis separately for older and younger siblings. The respiratory disease exposure index is the number of inpatient admissions with a respiratory disease primary diagnosis among children aged 13–71 months per 100 children in the focal child's municipality of birth during the first year of life, excluding any hospitalizations of an older sibling. Average labor market outcomes are calculated from siblings pairs at age 25-32. At each age, we require both of the siblings are observed. Income is reported in 2010 $USD. Income percentile is calculated among each year-age group. Test scores are converted into $z-$scores, which are standardized within each subject and test year. Test score data are only available for children born in 1986–2003. Average long-term health outcomes are calculated from siblings pairs at age 15-25. At each age, we also require both of the siblings are observed. Maternal educational attainment and parental marital/cohabiting status are measured at the time of childbirth, while household income is measured in the year before childbirth.

Table 2: Effect of Respiratory Disease Exposure Index on Respiratory Disease Hospitalizations in First Year of Life, Younger versus Older Siblings

	All Respiratory Hospitalizations in First Year of Life				
	(1)	(2)	(3)	(4)	(5)
Younger	0.039***		0.039***	0.007***	0.041***
	(0.002)		(0.002)	(0.002)	(0.003)
Disease index		0.017***	0.017***	0.010***	0.010***
		(0.001)	(0.001)	(0.001)	(0.001)
Younger x disease index				0.011***	0.012***
				(0.001)	(0.001)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	No	No	No	Yes
Observations	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982
Mean	0.068	0.068	0.068	0.068	0.068
25th to 75th pctile effect size				0.021	0.022

Notes: Each column in the table presents results from estimating different versions of model (1). The outcome is the number of hospitalizations with any respiratory disease primary diagnosis during the first year of the child's life. We report the coefficients on the indicator variable denoting the younger sibling ("Younger"), the respiratory disease exposure index ("Disease index"), and the interaction of these two variables. The respiratory disease exposure index is the number of inpatient admissions with any respiratory disease primary diagnosis among children aged 13–71 months per 100 children in each child's municipality of birth during the first year of life, excluding any hospitalizations of an older sibling. All specifications include municipality, year of birth, and month of birth fixed effects. Column (5) also includes the following family background controls: indicator for child gender, the sibling pair's birth spacing (in months) and the birth spacing interacted with the indicator for the younger child, mother's age and age squared, indicator for the mother being foreign-born, indicators for mother's education level (high school degree, college degree or higher), and an indicator for the parents being married or cohabiting at the time of childbirth. Standard errors are clustered on the child's municipality of birth in all models. The "25th to 75th pctile effect size" row reports the magnitude of the differential effect of an increase in the disease exposure index from the 25th to the 75th percentile of the distribution for younger siblings. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table 3: Effect of Respiratory Disease Exposure Index in First Year of Life on 9th Grade Danish Test Score, Younger versus Older Siblings

	9th Grade Danish Test Score			
	(1)	(2)	(3)	(4)
Younger	-0.140***		-0.140***	-0.118***
	(0.007)		(0.007)	(0.014)
Disease index		-0.002	-0.002	0.002
		(0.004)	(0.004)	(0.005)
Younger x disease index				-0.008*
				(0.005)
Municipality FEs	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes
Observations	469,170	469,170	469,170	469,170
Mean	0.100	0.100	0.100	0.100
25th to 75th pctile effect size				-0.013

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is the 9th grade Danish test score, which is converted into a $z-$score, standardized within each subject and test year. Test score data are only available for children born in 1986–2003. We require both of the siblings are observed in the data. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table 4: Effect of Respiratory Disease Exposure Index in First Year of Life on 9th Grade Math Test Score, Younger versus Older Siblings

	9th Grade Math Test Score			
	(1)	(2)	(3)	(4)
Younger	-0.146***		-0.146***	-0.135***
	(0.009)		(0.009)	(0.015)
Disease index		-0.000	-0.000	0.002
		(0.004)	(0.004)	(0.004)
Younger x disease index				-0.004
				(0.006)
Municipality FEs	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes
Observations	470,896	470,896	470,896	470,896
Mean	0.142	0.142	0.142	0.142
25th to 75th pctile effect size				-0.006

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is the 9th grade math test score, which is converted into a $z-$score, standardized within each subject and test year. Test score data are only available for children born in 1986–2003. We require both of the siblings are observed in the data. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

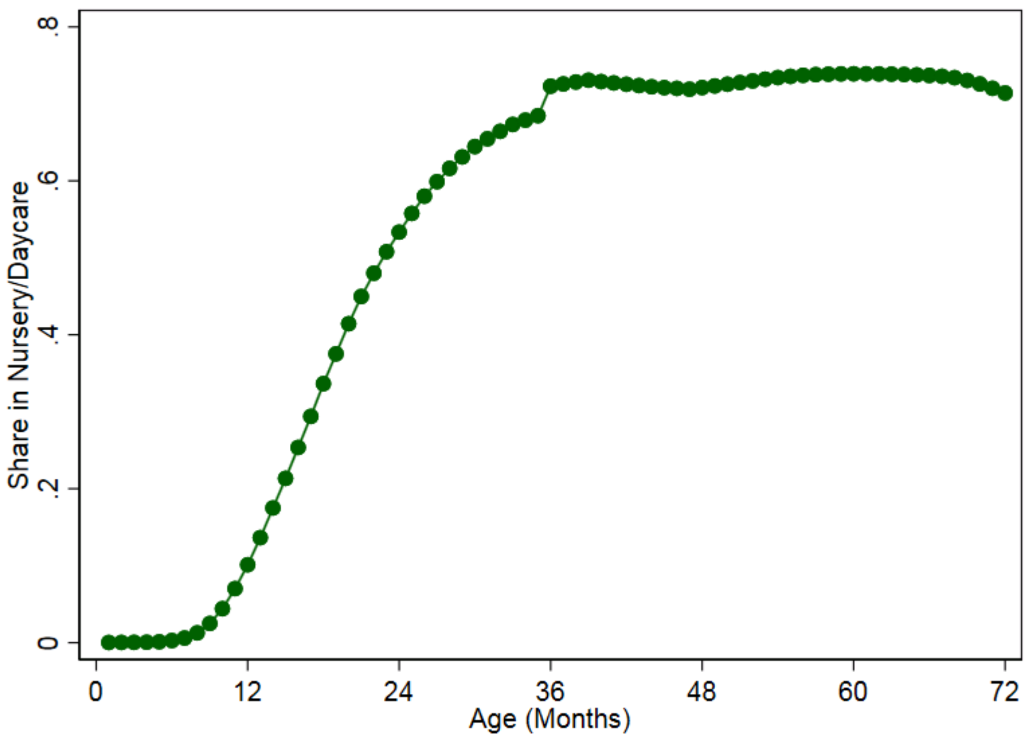
Table 5: Effect of Respiratory Disease Exposure Index in First Year of Life on Log Income (Conditional on Employment) at Ages 25–32, Younger versus Older Siblings

	Log Income at Age 25-32			
	(1)	(2)	(3)	(4)
Younger	-0.016***		-0.016***	-0.005
	(0.003)		(0.003)	(0.004)
Disease index		-0.000	-0.000	0.002
		(0.001)	(0.001)	(0.002)
Younger x disease index				-0.005***
				(0.001)
Municipality FEs	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes
Age FEs	Yes	Yes	Yes	Yes
Observations	1,613,376	1,613,376	1,613,376	1,613,376
Mean	10.923	10.923	10.923	10.923
25th to 75th pctile effect size				-0.007

Notes: See notes under Table 2 for more details about the specifications and variables. The sample includes sibling pairs at ages 25–32, with each observation at the person-by-age level. The outcome is the natural log of gross income (conditional on employed), converted into 2010 USD$. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

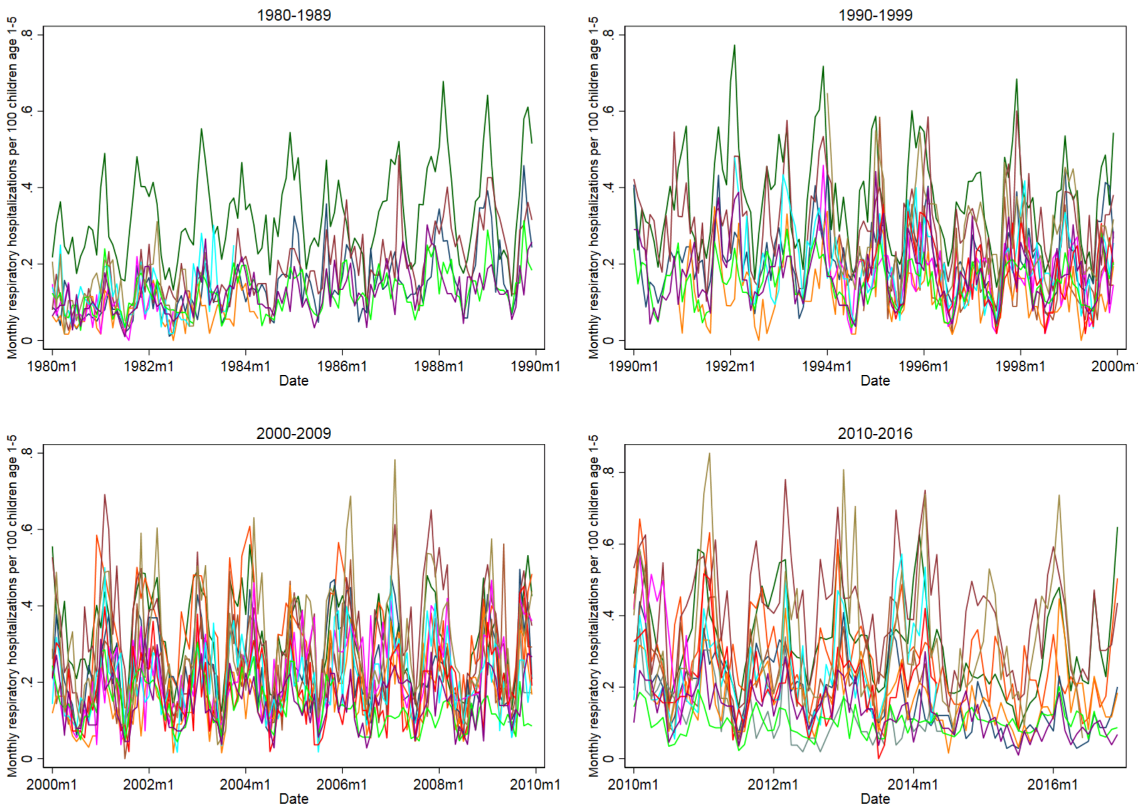
A Appendix Figures

Figure A1: Share of Children Attending Group Childcare by Child Age in Months



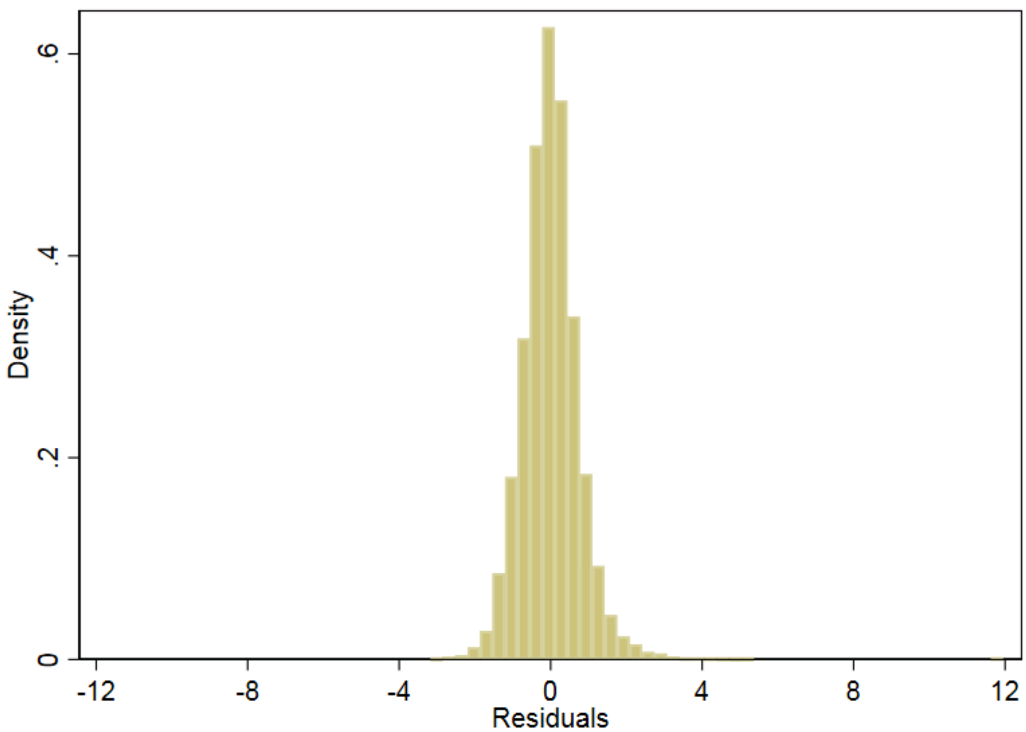
Notes: This graph shows the share of children who are attending childcare by age in months. We use data on enrollment in Danish childcare centers, which is reported annually in September of each year. This information is available to us over the period of September 1995 to September 2013.

Figure A2: Variation in the Respiratory Disease Index Over Time, 10 Largest Municipalities



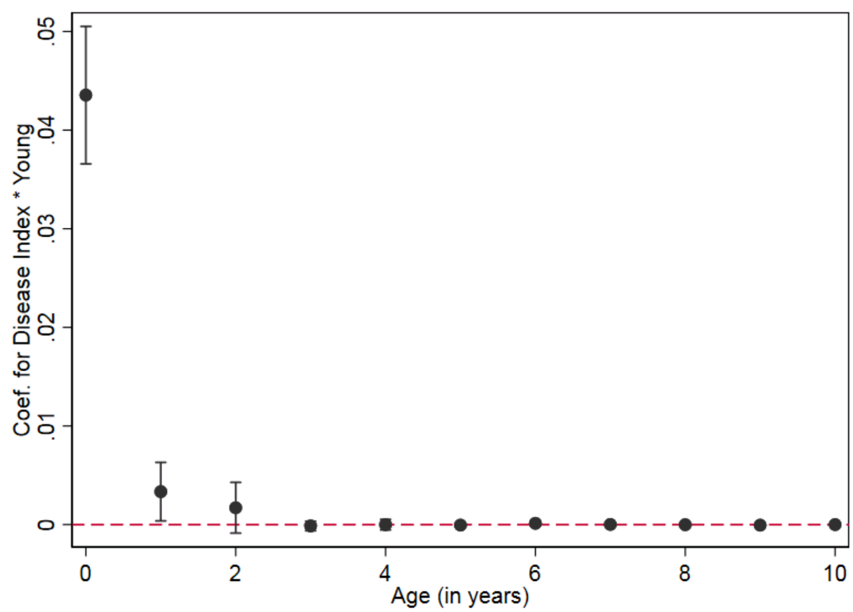
Notes: This figure shows the monthly variation in the respiratory disease index over time for each of the 10 largest municipalities (in terms of population size) in Denmark, separately for time periods of 1980-1989, 1990-1999, 2000-2009, and 2010-2016. The respiratory disease index refers to the number of respiratory disease hospitalizations per 100 children aged 13 to 71 months in each calendar year-month.

Figure A3: Distribution of Respiratory Disease Index Residuals from Municipality, Year, and Month Fixed Effects



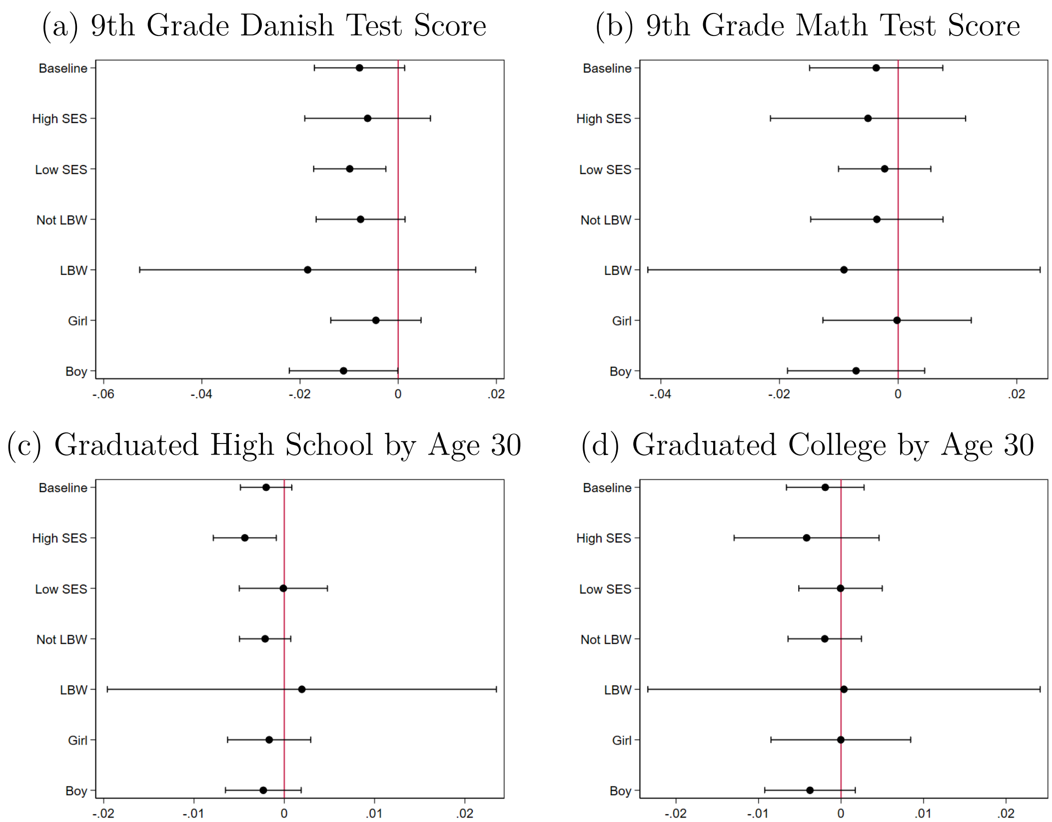
Notes: This histogram plots the residuals after regressing the respiratory disease index on municipality, year, and month fixed effects. The respiratory disease index refers to the number of respiratory disease hospitalizations per 100 children aged 13 to 71 months in each calendar year-month.

Figure A4: Effects of the RSV Index on the Annual Number of Younger Siblings' RSV Hospitalizations, by Age of Observation



Notes: This figure plots the coefficients and 95% confidence intervals on the interaction term between the RSV index and the younger sibling indicator from model (1), using as the outcome the annual number of hospitalizations with RSV diagnoses, measured at ages specified on the x-axis. At each age, we require both of the siblings are observed in the data. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 5 for more details about the definition of each subgroups and variables used in the specification. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

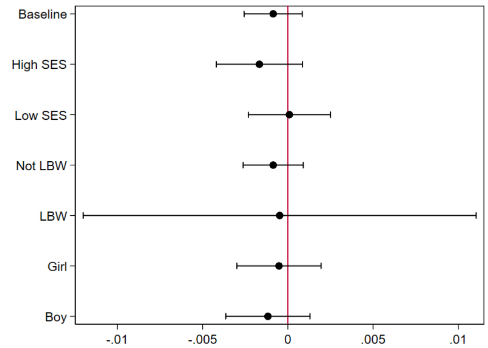
Figure A5: Heterogeneous Effects of the Respiratory Disease Exposure Index on Younger Siblings' Education Outcomes



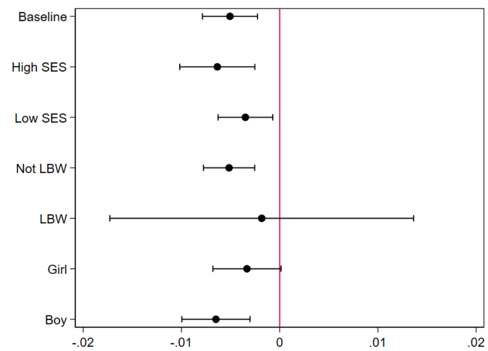
Notes: These figures plot the heterogeneity effects of the respiratory disease exposure on younger siblings education outcomes among different subgroups. The baseline coefficient and 95% confidence intervals are from the interaction term between the overall respiratory disease index and the younger sibling indicator from model (1). Effects by subgroups are from 3 separate regressions: 1) high vs. low socioeconomic status (SES); 2) low birth weight (LBW) status; and 3) child gender. In each regression, the full set of subgroup indicators are interacted with the younger sibling indicator, the disease index, and the younger sibling indicator × disease index interaction. Coefficients and 95% confidence intervals of the triple interaction term are plotted accordingly. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 4 for more details about the definition of each subgroups and variables used in the specification. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

Figure A6: Heterogeneous Effects of the Respiratory Disease Exposure Index on Younger Siblings' Labor Market Outcomes

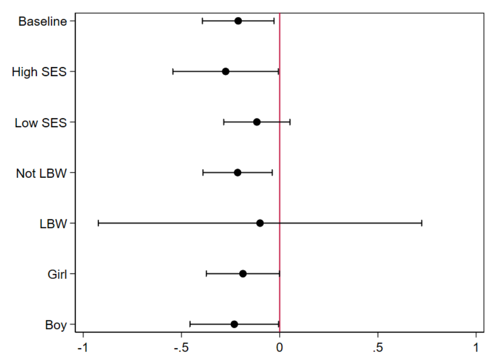
(a) Labor Force Participation



(b) Income (conditional on employed)



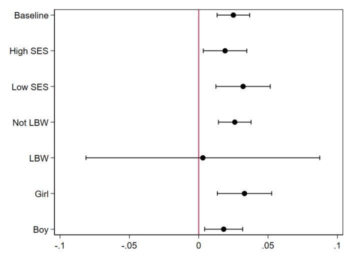
(c) Income Percentile



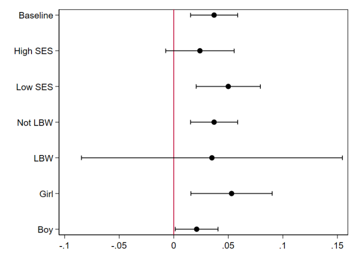
Notes: These figures plot the heterogeneity effects of the respiratory disease exposure on younger siblings labor market outcomes among different sub-populations. The baseline coefficient and 95% confidence intervals are from the interaction term between the overall respiratory disease index and the younger sibling indicator from model (1). Effects by subgroups are from 3 separate regressions: 1) high vs. low socioeconomic status (SES); 2) low birth weight (LBW) status; and 3) child gender. In each regression, the full set of sub-group indicators are interacted with the younger sibling indicator, the disease index, and the younger sibling indicator $\times$ disease index interaction. Coefficients and 95% confidence intervals of the triple interaction term are plotted accordingly. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 4 for more details about the definition of each subgroups and variables used in the specification. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

Figure A7: Heterogeneous Effects of the Respiratory Disease Exposure Index on Younger Siblings' Mental Health Care Outcomes

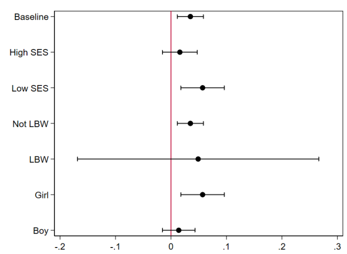
(a) Any Mental Health-Related Hospitalizations



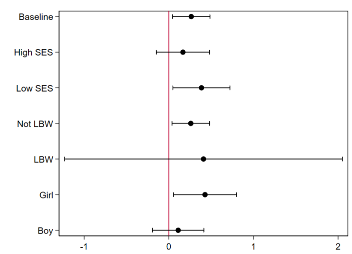
(b) Num. of Mental Health-Related Hospitalizations



(c) Any Psychiatrist Visits



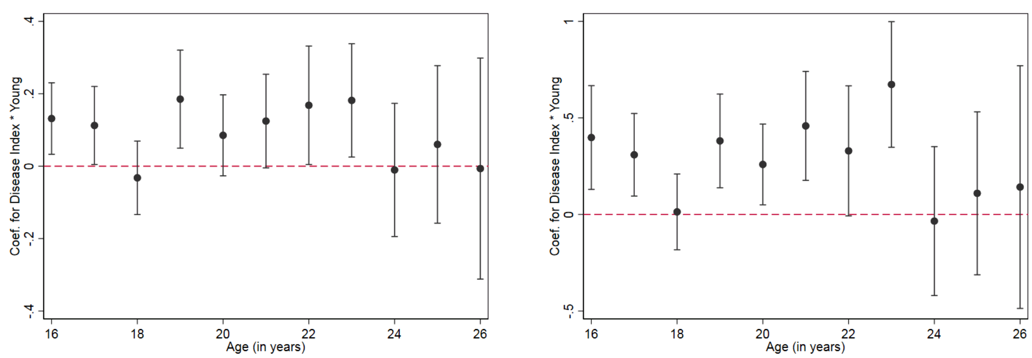
(d) Num. of Psychiatrist Visits



Notes: These figures plot the heterogeneity effects of the respiratory disease exposure on younger siblings mental health outcomes during ages 16-26 among different sub-populations. The sample includes sibling pairs at ages 16-26, with each observation at person-by-age level. The baseline coefficient and 95% confidence intervals are from the interaction term between the overall respiratory disease index and the younger sibling indicator from model (1) with age fixed effects. Effects by subgroups are from 3 separate regressions: 1) high vs. low socioeconomic status (SES); 2) low birth weight (LBW) status; and 3) child gender. In each regression, the full set of sub-group indicators are interacted with the younger sibling indicator, the disease index, and the younger sibling indicator × disease index interaction. Coefficients and 95% confidence intervals of the triple interaction term are plotted accordingly. All regressions include municipality, year of birth, month of birth, age fixed effects, and family background controls. See notes under Figure 4 for more details about the definition of each subgroups and variables used in the specification. Confidence intervals are constructed from two-way clustered standard errors at the individual and municipality of birth levels.

Figure A8: Effects of the Respiratory Disease Exposure Index on Younger Siblings' Hospitalizations for All Causes, by Age of Observation

(a) Any Hospitalization (b) Number of Hospitalizations



Notes: These figures plot the coefficients and 95% confidence intervals on the interaction term between the disease index and the younger sibling indicator from model (1), using the hospitalization outcomes measured at ages specified on the x-axes. At each age, we require both of the siblings are observed in the data. All regressions include municipality, year of birth, month of birth fixed effects, and family background controls. See notes under Figure 5 for more details about the specifications and variables. Confidence intervals are constructed from standard errors clustered on the child's municipality of birth.

B Appendix Tables

Table A1: Sample Construction Process

Sample Restriction	Observations
Birth cohort 1980-2015	2,221,433
Singleton first and second-born	1,373,056
Birth spacing gap at least 11 months	1,368,780
Drop sibling pairs with missing municipality of birth information, or born in municipalities with less than 1,000 children aged 13-71 months on average	1,335,548
Drop sibling pairs with missing parental control variables	1,163,982

Notes: This table shows how our sample size changes as we make various restrictions to arrive at our final analysis sample.

Table A2: Disease Exposure Index and Family Background Characteristics

	(1) Male	(2) Birth Spacing	(3) LBW	(4) VLBW	(5) Mother's Age	(6) Mother Foreign-Born	(7) Mother High School Graduated	(8) Mother College Graduated	(9) Parents Married/Cohabiting	(10) Father Log Income	(11) Mother Log Income	(12) Household Log Income	(13) Father Employed	(14) Mother Employed
Younger	-.002	-1.95*	-.0175***	-.00257***	2.78***	-.00693*	.00976	.0414***	.0174***	.135***	.165***	.132***	.052***	.0162
	(.00288)	(1.04)	(.00125)	(.000291)	(.0506)	(.00405)	(.00889)	(.00495)	(.00371)	(.026)	(.0177)	(.0196)	(.0147)	(.0152)
Disease index	-.00105	-.155	-.000468	-.0000908	-.139***	-.000031	.00314	-.0151**	-.00175	-.0141***	-.00914***	-.0124***	-.00674***	-.00122
	(.00079)	(.421)	(.000354)	(.000132)	(.034)	(.00149)	(.0021)	(.00615)	(.00128)	(.00411)	(.00311)	(.0037)	(.00237)	(.00277)
Younger x disease index	.000383	1.03**	.000743	.000148	.111***	.000894	-.00172	.000783	-.00329**	.00494	.0029	.00337	-.00192	-.00168
	(.000949)	(.422)	(.000469)	(.0000986)	(.0324)	(.00199)	(.00408)	(.00119)	(.00136)	(.00674)	(.0057)	(.00519)	(.0038)	(.00392)
Municipality FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982
Mean	0.514	41.961	0.033	0.004	28.569	0.045	0.770	0.335	0.944	10.899	10.605	11.512	0.870	0.754
25th to 75th pctile effect size	0.001	1.866	0.001	0.000	0.201	0.002	-0.003	0.001	-0.006	0.009	0.005	0.006	-0.003	-0.003

Notes: Each column in the table presents results from estimating model (1), separately for each of the dependent variables listed at the top. We report the coefficients on the indicator variable denoting the younger sibling ("Younger"), the respiratory disease exposure index ("Disease index"), and the interaction of these two variables. The disease exposure index is the number of inpatient admissions with a respiratory disease primary diagnosis among children aged 13–71 months per 100 children in each child's municipality of birth during the first year of life, excluding any hospitalizations of an older sibling. See notes under Table 2 for more details about the specifications. Standard errors are clustered on the child's municipality of birth in all models. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A3: Effect of Non-Infectious Digestive Disease Exposure Index on Non-Infectious Digestive Disease Hospitalizations in First Year of Life, Younger versus Older Siblings

	Non-infectious Digestive Disease Hospitalizations in First Year of Life (*1000)				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.243**		-0.242**	-0.175	-0.046
	(0.096)		(0.096)	(0.109)	(0.216)
Non-infectious digestive disease index		0.940	0.935	1.700	1.748
		(1.097)	(1.098)	(1.526)	(1.529)
Younger x Non-infectious disease index				-1.352	-1.371
				(1.426)	(1.411)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	No	No	No	Yes
Observations	687,914	687,914	687,914	687,914	687,914
Mean	0.968	0.968	0.968	0.968	0.968
25th to 75th pctile effect size				-0.093	-0.094

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is the number of hospitalizations with any non-infectious digestive disease primary diagnosis during the first year of the child's life (only available for children born after 1993). Standard errors are clustered on the child's municipality of birth in all models. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A4: Effect of Injury (incl. Poisoning) Exposure Index on Injury (incl. Poisoning) Hospitalizations in First Year of Life, Younger versus Older Siblings

	Injury (incl. Poisonings) Hospitalizations in First Year of Life (*1000)				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.264		-0.273	-1.279	-0.050
	(0.277)		(0.278)	(0.992)	(1.129)
Injury index		3.541***	3.544***	3.063***	3.020***
		(0.606)	(0.607)	(0.740)	(0.750)
Younger x injury index				0.871	0.891
				(0.928)	(0.940)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	No	No	No	Yes
Observations	687,914	687,914	687,914	687,914	687,914
Mean	7.326	7.326	7.326	7.326	7.326
25th to 75th pctile effect size				0.343	0.351

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is the number of hospitalizations with any injury (incl. poisoning) primary diagnosis during the first year of the child's life (only available for children born after 1993). Standard errors are clustered on the child's municipality of birth in all models. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A5: Effect of RSV Index on RSV Hospitalizations in the First Year of Life, Younger versus Older Siblings

	RSV Hospitalizations in First Year of Life				
	(1)	(2)	(3)	(4)	(5)
Younger	0.018***		0.018***	0.014***	0.029***
	(0.001)		(0.001)	(0.001)	(0.001)
RSV index		0.040***	0.040***	0.016***	0.017***
		(0.004)	(0.004)	(0.003)	(0.003)
Younger x RSV index				0.045***	0.044***
				(0.003)	(0.004)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	No	No	No	Yes
Observations	687,914	687,914	687,914	687,914	687,914
Mean	0.018	0.018	0.018	0.018	0.018
25th to 75th pctile effect size				0.005	0.005

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is the number of hospitalizations with an RSV primary diagnosis during the first year of the child's life (only available for children born after 1993). The disease index is constructed using hospitalizations for RSV only (rather than all hospitalizations for respiratory conditions). Standard errors are clustered on the child's municipality of birth in all models. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A6: Robustness of Results on Respiratory Disease Hospitalizations in First Year of Life

	Respiratory Disease Hospitalizations in First Year of Life				
	(1)	(2)	(3)	(4)	(5)
Younger	0.041***	0.040***	0.037***	0.039***	0.040***
	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)
Disease index	0.010***	0.007***	0.004***		
	(0.001)	(0.001)	(0.002)		
Younger x disease index	0.012***	0.012***	0.013***		
	(0.001)	(0.001)	(0.001)		
Disease index (# Diagnosis)				0.005***	
				(0.001)	
Younger x disease index (# Diagnosis)				0.009***	
				(0.000)	
Disease index (# Kids)					0.011***
					(0.001)
Younger x disease index (# Kids)					0.013***
					(0.001)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Observations	1,163,982	1,163,982	1,163,982	1,163,982	1,163,982
Mean	0.068	0.068	0.068	0.068	0.068
25th to 75th pctile effect size	0.022	0.022	0.023	0.022	0.022

Notes: Each column in the table presents results from estimating different versions of model (1). The outcome is the number of hospitalizations with a respiratory disease primary diagnosis. Column (1) presents results using the baseline model. Column (2) adds municipality-specific linear time trends, while column (3) adds maternal fixed effects. Column (4) uses a disease index in which we count number of diagnoses for respiratory conditions in hospitalizations including both primary and non-primary diagnoses. Column (5) uses a disease index in which we calculate the number of children with at least one respiratory disease diagnosis (i.e., counting the number of children and not the total number of diagnoses). See notes under Table 2 for more details about our baseline model and control variables. Standard errors are clustered on the child's municipality of birth in all models. The "25th to 75th pctile effect size" row reports the magnitude of the differential effect of an increase in the disease exposure index from the 25th to the 75th percentile of the distribution for younger siblings. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A7: Robustness of Results on 9th Grade Danish Test Score

	9th Grade Danish Test Score				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.118***	-0.129***	-0.115***	-0.123***	-0.117***
	(0.014)	(0.013)	(0.013)	(0.013)	(0.014)
Disease index	0.002	-0.003	0.001		
	(0.005)	(0.006)	(0.006)		
Younger x disease index	-0.008*	-0.005	-0.009**		
	(0.005)	(0.005)	(0.004)		
Disease index (# Diagnosis)				-0.000	
				(0.004)	
Younger x disease index (# Diagnosis)				-0.004	
				(0.003)	
Disease index (# Kids)					0.003
					(0.006)
Younger x disease index (# Kids)					-0.009*
					(0.005)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Observations	469,170	469,170	469,170	469,170	469,170
Mean	0.100	0.100	0.100	0.100	0.100
25th to 75th pctile effect size	-0.013	-0.008	-0.015	-0.011	-0.014

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The outcome is the 9th grade Danish test score, which is converted into a $z-$score, standardized within each subject and test year. Test score data are only available for children born in 1986–2003. We require both of the siblings are observed in the data. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A8: Robustness of Results on 9th Grade Math Test Score

	9th Grade Math Test Score				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.135***	-0.123***	-0.136***	-0.141***	-0.135***
	(0.015)	(0.020)	(0.013)	(0.014)	(0.015)
Disease index	0.002	0.009*	0.008*		
	(0.004)	(0.005)	(0.005)		
Younger x disease index	-0.004	-0.008	-0.003		
	(0.006)	(0.008)	(0.004)		
Disease index (# Diagnosis)				0.001	
				(0.002)	
Younger x disease index (# Diagnosis)				-0.001	
				(0.004)	
Disease index (# Kids)					0.001
					(0.004)
Younger x disease index (# Kids)					-0.004
					(0.006)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Observations	470,896	470,896	470,896	470,896	470,896
Mean	0.142	0.142	0.142	0.142	0.142
25th to 75th pctile effect size	-0.006	-0.013	-0.005	-0.003	-0.007

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The outcome is the 9th grade math test score, which is converted into a $z-$score, standardized within each subject and test year. Test score data are only available for children born in 1986–2003. We require both of the siblings are observed in the data. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A9: Robustness of Results on High School Graduation by Age 30

	Graduated High School by Age 30				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.016***	-0.020***	-0.008	-0.015**	-0.016***
	(0.006)	(0.007)	(0.008)	(0.006)	(0.006)
Disease index	0.005**	0.005*	0.010*		
	(0.003)	(0.003)	(0.005)		
Younger x disease index	-0.002	-0.000	-0.005*		
	(0.001)	(0.002)	(0.003)		
Disease index (# Diagnosis)				0.005***	
				(0.002)	
Younger x disease index (# Diagnosis)				-0.002*	
				(0.001)	
Disease index (# Kids)					0.006*
					(0.003)
Younger x disease index (# Kids)					-0.002
					(0.002)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Observations	148,288	148,288	148,288	148,288	148,288
Mean	0.844	0.844	0.844	0.844	0.844
25th to 75th pctile effect size	-0.002	-0.000	-0.005	-0.003	-0.002

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The outcome is an indicator for graduating high school by age 30. We require both of the siblings are observed in the data. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A10: Robustness of Results on College Graduation by Age 30

	Graduated College by Age 30				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.045***	-0.046***	-0.037***	-0.045***	-0.045***
	(0.009)	(0.010)	(0.011)	(0.008)	(0.009)
Disease index	0.002	0.002	0.005		
	(0.003)	(0.004)	(0.005)		
Younger x disease index	-0.002	-0.002	-0.003		
	(0.002)	(0.003)	(0.003)		
Disease index (# Diagnosis)				0.003	
				(0.002)	
Younger x disease index (# Diagnosis)				-0.002	
				(0.002)	
Disease index (# Kids)					0.001
					(0.004)
Younger x disease index (# Kids)					-0.002
					(0.003)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Observations	148,288	148,288	148,288	148,288	148,288
Mean	0.437	0.437	0.437	0.437	0.437
25th to 75th pctile effect size	-0.002	-0.002	-0.004	-0.003	-0.002

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The outcome is an indicator for graduating college by age 30. We require both of the siblings are observed in the data. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A11: Robustness of Results on Labor Force Participation at Ages 25–32

	Labor Force Participation at Age 25-32				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.012***	-0.012***	-0.003	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Disease index	0.005***	0.003*	0.004**		
	(0.002)	(0.002)	(0.002)		
Younger x disease index	-0.001	-0.001	-0.001		
	(0.001)	(0.001)	(0.001)		
Disease index (# Diagnosis)				0.004***	
				(0.001)	
Younger x disease index (# Diagnosis)				-0.001	
				(0.001)	
Disease index (# Kids)					0.005**
					(0.002)
Younger x disease index (# Kids)					-0.001
					(0.001)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	2,391,872	2,391,872	2,372,143	2,391,872	2,391,872
Mean	0.693	0.693	0.695	0.693	0.693
25th to 75th pctile effect size	-0.001	-0.001	-0.001	-0.001	-0.001

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 25–32, with each observation at the person-by-age level. The outcome is an indicator for being in the labor force. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A12: Robustness of Results on Log Income (Conditional on Employed) at Ages 25–32

	Log Income at Age 25-32				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.005	-0.010***	0.005	-0.005	-0.005
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Disease index	0.002	0.000	0.006***		
	(0.002)	(0.001)	(0.002)		
Younger x disease index	-0.005***	-0.003***	-0.008***		
	(0.001)	(0.001)	(0.001)		
Disease index (# Diagnosis)				0.002	
				(0.001)	
Younger x disease index (# Diagnosis)				-0.003***	
				(0.001)	
Disease index (# Kids)					0.002
					(0.002)
Younger x disease index (# Kids)					-0.005***
					(0.002)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	1,613,376	1,613,376	1,589,702	1,613,376	1,613,376
Mean	10.923	10.923	10.927	10.923	10.923
25th to 75th pctile effect size	-0.007	-0.004	-0.011	-0.007	-0.007

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 25–32, with each observation at the person-by-age level. The outcome is the natural log of gross income (conditional on employed), converted into 2010 USD$. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A13: Robustness of Results on Income Percentile at Ages 25–32

	Income Percentile at Age 25-32				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.805***	-0.961***	-0.086	-0.792***	-0.804***
	(0.215)	(0.201)	(0.212)	(0.218)	(0.218)
Disease index	0.266**	0.095	0.412***		
	(0.123)	(0.102)	(0.116)		
Younger x disease index	-0.211**	-0.149*	-0.301***		
	(0.093)	(0.079)	(0.081)		
Disease index (# Diagnosis)				0.221**	
				(0.095)	
Younger x disease index (# Diagnosis)				-0.157**	
				(0.067)	
Disease index (# Kids)					0.267**
					(0.133)
Younger x disease index (# Kids)					-0.222**
					(0.100)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	2,391,872	2,391,872	2,372,143	2,391,872	2,391,872
Mean	56.337	56.337	56.370	56.337	56.337
25th to 75th pctile effect size	-0.306	-0.215	-0.433	-0.311	-0.306

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 25–32, with each observation at the person-by-age level. The outcome is the income percentile (calculated using the population of the same age in each year). Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A14: Robustness of Results on Any Mental Health-Related Hospitalizations Annually at Ages 16–26

	Any Mental-Related Hospitalizations Annually 16-26 * 100				
	(1)	(2)	(3)	(4)	(5)
Younger	0.035*	0.058***	-0.022	0.031	0.033*
	(0.019)	(0.019)	(0.022)	(0.020)	(0.019)
Disease index	0.020*	0.010	-0.000		
	(0.011)	(0.009)	(0.014)		
Younger x disease index	0.025***	0.017***	0.042***		
	(0.006)	(0.006)	(0.007)		
Disease index (# Diagnosis)				0.009	
				(0.008)	
Younger x disease index (# Diagnosis)				0.019***	
				(0.004)	
Disease index (# Kids)					0.021*
					(0.012)
Younger x disease index (# Kids)					0.027***
					(0.006)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	5,204,048	5,204,048	5,184,713	5,204,048	5,204,048
Mean	0.405	0.405	0.406	0.405	0.405
25th to 75th pctile effect size	0.039	0.027	0.066	0.042	0.041

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 16–26, with each observation at the person-by-age level. The outcome is an indicator for having at least one mental health-related hospitalization during the observed age. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A15: Robustness of Results on the Number of Mental Health-Related Hospitalizations Annually at Ages 16–26

	Number of Mental-Related Hospitalizations Annually 16-26 * 100				
	(1)	(2)	(3)	(4)	(5)
Younger	0.041	0.070**	-0.037	0.037	0.036
	(0.029)	(0.028)	(0.036)	(0.029)	(0.030)
Disease index	0.034**	0.014	0.007		
	(0.014)	(0.013)	(0.026)		
Younger x disease index	0.037***	0.026***	0.062***		
	(0.011)	(0.010)	(0.015)		
Disease index (# Diagnosis)				0.016*	
				(0.009)	
Younger x disease index (# Diagnosis)				0.028***	
				(0.008)	
Disease index (# Kids)					0.037**
					(0.015)
Younger x disease index (# Kids)					0.040***
					(0.012)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	5,204,048	5,204,048	5,184,713	5,204,048	5,204,048
Mean	0.484	0.484	0.484	0.484	0.484
25th to 75th pctile effect size	0.058	0.041	0.098	0.061	0.061

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 16–26, with each observation at the person-by-age level. The outcome is the number of mental health-related hospitalizations during the observed age. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A16: Robustness of Results on Any Psychiatrist Visits Annually at Ages 16–26

	Any Psychiatrist Visit Annually 16-26 * 100				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.106***	-0.087**	-0.166***	-0.116***	-0.114***
	(0.040)	(0.040)	(0.049)	(0.039)	(0.040)
Disease index	-0.005	0.017	-0.023		
	(0.022)	(0.025)	(0.027)		
Younger x disease index	0.035***	0.029**	0.058***		
	(0.012)	(0.012)	(0.016)		
Disease index (# Diagnosis)				-0.013	
				(0.015)	
Younger x disease index (# Diagnosis)				0.028***	
				(0.009)	
Disease index (# Kids)					-0.009
					(0.024)
Younger x disease index (# Kids)					0.040***
					(0.013)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	4,779,471	4,779,471	4,760,527	4,779,471	4,779,471
Mean	1.093	1.093	1.095	1.093	1.093
25th to 75th pctile effect size	0.055	0.045	0.090	0.062	0.059

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 16–26, with each observation at the person-by-age level. The outcome is an indicator for visiting the psychiatrist for at least once during the observed age. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A17: Robustness of Results on the Number of Psychiatrist Visits Annually at Ages 16-26

	Number of Psychiatrist Visits Annually 16-26 * 1000				
	(1)	(2)	(3)	(4)	(5)
Younger	-1.091***	-0.758*	-1.588***	-1.201***	-1.164***
	(0.393)	(0.428)	(0.434)	(0.378)	(0.394)
Disease index	-0.059	0.141	-0.210		
	(0.204)	(0.223)	(0.243)		
Younger x disease index	0.265**	0.148	0.491***		
	(0.113)	(0.126)	(0.105)		
Disease index (# Diagnosis)				-0.118	
				(0.138)	
Younger x disease index (# Diagnosis)				0.225***	
				(0.081)	
Disease index (# Kids)					-0.074
					(0.220)
Younger x disease index (# Kids)					0.309**
					(0.122)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	Yes	No	No	No
Mother FEs	No	No	Yes	No	No
Age FEs	Yes	Yes	Yes	Yes	Yes
Observations	4,779,471	4,779,471	4,760,527	4,779,471	4,779,471
Mean	7.658	7.658	7.676	7.658	7.658
25th to 75th pctile effect size	0.413	0.231	0.766	0.485	0.457

Notes: See notes under Appendix Table A6 for more details about the specifications and variables. The sample includes sibling pairs at ages 16–26, with each observation at person-by-age level. The outcome is the number of psychiatrist visits during the observed age. Age fixed effects are included in all regressions. Standard errors are clustered on the individual and municipality of birth level. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.