Causes and Consequences of Early Childhood Infectious Disease *

N. Meltem Daysal† Hui Ding‡

Maya Rossin-Slater§ Hannes Schwandt¶

February 15, 2021

Preliminary and incomplete! Please do not circulate.

Abstract

Infectious diseases pose an important public health concern, and can generate large disruptions in the economy and societal functioning more broadly. While large-scale global pandemics such as the Spanish Influenza of 1918 or the novel Corona virus receive substantial attention due to their dramatic health and social impacts, more common infectious diseases that affect populations on an annual basis can also have important aggregate costs because of the large share of the population that is affected on a regular basis. Yet relatively little is known about the causes and consequences of infectious diseases, especially over a long-term horizon. This paper explores the causes and consequences of infectious diseases, focusing on a widely recognized "disease hub" in the population: families with young children. In the first part of the paper, we analyze sibling pairs in Danish register data and document substantially higher rates of respiratory diseases during the first year of life for second-born children than their first-born counterparts during the same stage of life. The patterns in the data suggest the older sibling brings the diseases home which then infect the younger sibling. In the second part of the paper, we construct a disease index at the municipality level to obtain exogenous variation in infectious disease exposure which allows to measure causal impacts on long-term outcomes. We find that higher rates of infectious diseases during infancy have negative impacts on educational and labor market outcomes in young adulthood.

*We thank Jon Guryan and participants at various seminars for helpful comments.

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

‡Department of Economics, Stanford University. Email: khding@stanford.edu.

§Department of Medicine, 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

Infectious diseases pose an important public health risk and can generate large disruptions in the economy and societal functioning more broadly. While global pandemics such as the novel Corona virus or the Spanish Influenza of 1918 receive substantial attention due to their dramatic health and social impacts, more common infectious diseases that affect populations on an annual basis can also have important aggregate costs because of the large share of the population that is affected on a regular basis. In this paper, we focus on a population that is particular exposed and vulnerable to common infectious diseases: Infants during their first year of life. We use rich administrative data from Denmark to analyze both the mechanisms through which infants become infected by respiratory diseases as well as the consequences of this early disease exposure for later life development. Infections early in life might be particularly costly if they have lasting developmental impacts. While a large economic literature has shown long-term impacts of in-utero exposure to severe shocks, such as as famines or natural disasters, relatively less is known about the infancy period and the role of more moderate shocks such as common infectious diseases (Currie and Almond, 2011; Almond et al., 2018). What makes infectious diseases particularly relevant is the fact that there are many known practices and policies, ranging from mask-wearing and vaccination to social distancing and lockdowns, to reduce disease spread. But such measures can be painful and costly, as the current pandemic has once more shown. Optimizing behavioral and policy responses to disease spread therefore requires an understanding of how diseases reach vulnerable populations, such as infants, and how they impact those who get infected.

In the first part of this paper, we provide descriptive evidence for all first- and second-born siblings born in Denmark between 1980 and 2015 suggesting that older siblings bringing home seasonal disease outbreaks constitute a primary mechanism how infectious respiratory diseases reach younger siblings. In particular, we show that inpatient hospitalizations for respiratory illness during the first year of life are up to 300% higher among second-born children than their first-born counterparts during the same stage of life. Moreover, the difference is larger if the younger sibling is born in the fall or winter (and hence exposed to the "winter wave" during the first months of life) and if the birth spacing between siblings is smaller. While these patterns suggest a causal pathway based on family structure, we do not use this variation to analyze long-term effects of infectious disease because family structure is endogenous to other determinants of health and human capital.

In the second part of the paper, we introduce a respiratory sickness index at the municipality level as an exogenous source of disease spread that is not driven by the structure of a given family or by other family-specific determinants of health and human capital. The disease index is composed of respiratory hospitalizations among all children in a municipality from age 1 through age 6 and excludes the focal infant's older sibling. We then link our constructed disease indices to children's actual observed sickness in infancy as well as to their human capital and labor market outcomes during young adulthood. We find that local sickness indices have strong impacts on infant health for both siblings but that the effects are substantially larger for the younger sibling. In the long run, these exogenous shocks to infant health are associated with worse educational and labor market outcomes, again with substantially stronger effects for second-born children.

We also analyze the causal impacts of infancy infections on respiratory disease hospitalizations later in childhood. Higher respiratory disease hospitalization rates during the first year of life, driven by local disease outbreaks that are brought home by an older sibling, lead to lower hospitalization rates for the same diseases at age three and four. This finding is in line with positive impacts on immunity and resistance built up via early disease exposure.

This project contributes to a growing literature analyzing moderate but common shocks to child development (Almond et al., 2018). Moreover, we zoom in on the first year of life, a developmental period that has received relatively less attention compared to the in-utero period and later childhood years. Our analysis also provides new mechanisms for sibling spillovers that have been extensively studied in the context of parental investments in child quality (Black et al. 2020, Daysal et al. 2020). Finally, our analysis points to family structure as an important social driver of disease spread and uncovers detrimental consequences of infectious diseases for child development (Adda 2016, Schwandt 2019).

Our findings are also informative to assess the costs of the current pandemic for young children. While children and infants are generally not strongly affected by the Coronavirus, the biggest impact of the pandemic might run through the reduction of other common infectious diseases such as the flu which spread has been limited by social distancing and lockdowns.

School closures in particular have limited the ability of viruses to spread to infants via older siblings bringing them home from schools. As a result, the pandemic might have positive long-term effects on the "COVID-19 generation" that was born right before and during the pandemic. At the same time, our results suggest that the low disease environment for infants during the pandemic might be followed by increased rates of infectious disease hospitalizations in the subsequent years when children become exposed to diseases without immunity developed during infancy.

2 Data and Sample

We use population register data from Denmark for the years 1980 to 2016. 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.

Outcome variables. As outcomes, we study hospitalizations during childhood, as well as measures of long-term educational achievement and economic well-being.

More specifically, our key short-run outcome is the number of hospitalizations with a respiratory disease diagnosis during the first year of life. We measure this outcome using the *National Patient Register*, which includes all inpatient admissions to public and private hospitals, along with International Classification of Disease (ICD) diagnosis and procedure codes (Lynge et al., 2011). We classify inpatient visits with the following primary diagnosis codes as respiratory disease-related: ICD-9 codes starting with "46," "47," "48," "490," "079," and "783"; ICD-10 codes starting with "B974" or "J" (excluding "J4").

For educational outcomes, we consider 9th grade test scores and educational attainment by age 24. Test scores come from the *Academic Achievement Register*, which exists from 2001 onward. We consider Danish (reading) and mathematics test scores separately, and standardize the scores within subject and birth cohort such that they have a mean of zero and a standard deviation of one. Educational attainment comes from the *Education Register*, which contains the highest level of completed schooling from administrative school records. We create indicators for having a high school or a college degree by age 24.

Finally, we use two registers to measure labor market outcomes at age 26. We use the

Register-Based Labour Force Statistics to characterize employment status. This data set is based on tax records, and records the employment 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 employed and zero otherwise (i.e., those who are unemployed and those who are out of the labor force are both coded as zero). We also use the *Income Statistics Register* to calculate the natural log of gross personal income at age 26, converted into 2010 $USD.[1]

Control variables. We observe a rich set of child and parent characteristics, using the previously defined 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.[2]

We include the following variables as controls in some of our specifications: child gender, maternal age at the time of childbirth, indicators for the mother being foreign-born, maternal education level in the year before childbirth (less than high school, high school degree or higher, college degree or higher), and parental marital/cohabitation status in the year of childbirth.

Analysis sample. To construct our analysis sample, we begin with the universe of 2,221,433 children born between 1980 and 2015 in Denmark with non-missing information on municipality of birth, 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 born in municipalities that have an average of at least 1,000 children aged 13–71 months

[1]The *Income Statistics Register* is based on tax records for all Danish residents aged 14 and older, and includes more than 160 variables measured on an annual basis such as salary income, taxes, public transfer payments, capital income, and private pension contributions (Baadsgaard and Quitzau, 2011).

[2]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 **XXX** percent of observations in our analysis period.

over the sample period, which ensures that we have sufficient observations to calculate the respiratory disease exposure index as described in Section 3 below.[3]

These restrictions leave us with a sample of 1,335,548 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 limit to children born between 1986 and 2000 because test score data begin in 2001 and we need to observe children when they are in 9th grade (around age 16). To study educational attainment by age 24, we use children born between 1980 and 1993, while to study labor market outcomes at age 26, we use children born between 1980 and 1990.

3 Descriptive Analysis and Empirical Design

3.1 Differences in Respiratory Disease Hospitalizations between Older and Younger Siblings

Our goal is to analyze the causal effects of infectious disease exposure in early childhood on health, educational, and economic outcomes. To motivate our empirical design, 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 infectious disease is 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.

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 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 fact

[3] This restriction leads to drop children born in 8 small municipalities.

that the vast majority of Danish children are at home with their mothers during their first year of life, and begin attending group childcare after they turn one year old.[4] Thus, after age one, younger and older siblings are similarly likely to be exposed to infectious viruses in group care environments, whereas only second-born children can have viruses be "brought home" to them by their older siblings before they turn one.

In Figure 2, we explore the role of infectious 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 infectious disease outbreaks (such as influenza) are more prevalent—children born in November, December, and January have highest hospitalization rates in the first three months of life; those born in August, September, and October have highest hospitalization rates at 3 to 6 months old; those born in May, June, and July have highest hospitalization rates at 7 to 9 months old; and those born in February, March, and April have 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 to three months old, suggesting that they are particularly susceptible to severe respiratory infections during early infancy.

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

[4] In Denmark, new parents can receive up to 52 weeks of parental leave with partial pay.

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 disease spreads within the family because older children "bring home" infectious diseases that they pick up in their local community (e.g., their childcare center). This analysis informs our empirical strategy for estimating the causal effects of early childhood 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 and 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 use the *National Patient Register* data to construct our disease exposure index as follows. First, for each municipality and in each calendar year-month, we calculate the number of respiratory disease hospitalizations per 100 children aged 13 to 71 months.[5] Second, for each child in our sibling analysis sample, we assign the monthly hospitalization rate that we calculated 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 disease exposure index on younger versus older siblings. Specifically, our regression models take the form:

[5]We use the 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 school at age 6 in Denmark.

$$Y^a_{itm} = \beta_0 + \beta_1 Younger_i + \beta_2 Index_{itm} + \beta_3 Younger_i \times Index_{itm} + \mu_m + \theta_t + \gamma' X_i + \epsilon_{itm} \quad (1)$$

for each child i born in year-month t in municipality m. Y^a_{itm} is an outcome measured at age a. $Younger_i$ is an indicator set to 1 for younger siblings, and captures birth order effects on our outcomes of interest. $Index_{itm}$ 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 are year and month of birth fixed effects that control for cohort 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, mother's age and age squared, indicators for mother's education level (less than high school, high school degree, college degree or higher), and an indicator for parents being married or cohabiting. We cluster standard errors on 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 later 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 generally robust across these models. Second, we investigate several types of "balance tests". In particular, Appendix Table A1 presents results from estimating different versions of model (1), using as the outcome each child's *predicted* number of respiratory disease hospitalizations during the first year of life, which is predicted after regressing the child's actual number of hospitalizations on the vector of individual and family background control variables, X_i. Column (1) through (3) show that while there are no significant differences in the predicted outcome by birth order, there is a positive correlation between the respiratory disease exposure index and the predicted outcome overall (i.e., the main effect of the disease index is positive and statistically significant). However, our key coefficient on the interaction between the younger child indicator and the disease index is not statistically significant, and this is the case in models both without and with municipality-specific linear time trends (columns 4 and 5, respectively).

In addition, we estimate model (1) using several individual background characteristics as outcomes. As shown in Appendix Table XXX, our key coefficient on the interaction term between the younger sibling indicator and the disease exposure index is insignificant in all of the models.

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 first panel shows that mothers are on average aged 26.6 years at the time of their first birth and 30.2 years at the time of their second birth. Approximately seven percent of mothers in our sample are foreign-born. About 73.9 and 77.2 percent of mothers have a high school degree at the time of the first and second birth, respectively, while 8.8 and 10.3 percent have a college degree, respectively. Approximately 80 percent of parents are married or cohabiting at the time of the first birth, while 91.6 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.

The second panel highlights some differences in child outcomes by birth order. Compared to older siblings, younger siblings have higher average birth weight (3575.7 versus 3419.9 grams for younger versus older siblings, respectively). 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 and similar local exposure to respiratory disease, younger siblings are nearly twice more likely to be hospitalized for respiratory conditions during their first year of life (the hospitalization rates are 0.09 and 0.05 for younger and older siblings, respectively). Moreover, consistent with prior literature on the impacts of birth order (e.g., **?**), younger siblings have lower academic and economic outcomes than their older counterparts.

4 Results

Effects of local respiratory disease exposure on children's hospitalizations. Table 2 presents results from estimating equation (1) using as the outcome the number of hospitalizations with a respiratory disease primary diagnosis during the first year of a child's life. We report the coefficients on the indicator denoting the younger sibling, the respiratory disease exposure index, and the interaction of these two variables. Column (1) shows that, consistent with the evidence in Figures 1 through 3, younger siblings are 4.3 percentage points (63.2 percent relative to the sample mean) more likely to be hospitalized 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 sample overall, 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) through (6), we find that there is a significant differential effect of local respiratory disease exposure on younger siblings relative to older siblings. We find that an additional respiratory hospitalization among slightly older children in the local community per 100 children has a 1.2 percentage point higher impact on a younger child's own number of hospitalizations in the first year of life than on their older sibling's number of hospitalizations. This relationship is extremely robust across specifications that include family background control variables and municipality-specific trends (columns 5 and 6, 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 2.2 percentage point 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.

Effects of local respiratory disease exposure before age one on long-term 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 long-term educational and economic outcomes. Tables 3 and 4 present results using as outcomes standardized 9th grade Danish and math scores, respectively. Columns (1)-(3) demonstrate ("main") birth order and early life disease exposure effects: younger children have lower test scores than older children, and a higher disease index in the first year of life is associated with a lower test score on average. However, younger children experience a differential penalty from local disease exposure in early life. An additional respiratory hospitalization among slightly older children in the municipality per 100 children reduces 9th Danish and math test scores each by about 0.014 of a standard deviation more for younger siblings than older siblings. The 25th to 75th percentile increase in the disease index amounts to an additional 0.024 of a standard deviation penalty on both test scores for the younger siblings relative to the older siblings. The results are robust to the inclusion of municipality-specific trends in column (5).

Tables 5 and 6 present results for high school and college graduation by age 24, respectively. As with test scores, we find an overall birth order effect—younger siblings are on average 8.3 and 2.8 percentage points less likely to graduate high school and college, respectively, than their older counterparts. However, we no longer observe an overall effect of early life disease exposure on these outcomes—the coefficients on the disease exposure index in columns (2) and (3) are very small in magnitude and insignificant. At the same time, columns (4) and (5) show that the younger siblings do experience an adverse effect of early life disease exposure on educational attainment relative to older siblings. In particular, column (4) shows that moving from the 25th to the 75th percentile in the respiratory disease exposure index distribution is associated with an additional 0.7 and 0.6 percentage point reduction in the likelihood high school and college graduation, respectively, for younger siblings relative to older siblings. These results are again robust to including municipality-specific trends (column 5).

Finally, Tables 7 and 8 report results for labor force participation and log income at age

26. We do not find a significant overall effect or a differential effect by birth order of local disease exposure in early childhood on the likelihood of being in the labor force. However, we do observe significant impacts on age-26 income. Table 8 shows that while for older siblings, there is a positive relationship between the disease exposure index and income, younger siblings experience a differential penalty. Column (4) demonstrates that moving from the 25th to the 75th percentile in the respiratory disease exposure index distribution is associated with an additional 1.8 percent reduction in age-26 income relative to older siblings. As with all our estimates, these results are robust to municipality-specific trends in column (5).

Differential effects by age of observation. Panel (a) of Figure 4 plots the coefficients and 95% confidence intervals on our key variable—the interaction between the younger sibling indicator and the disease exposure index—from separate models that use as outcomes the annual number of respiratory disease hospitalizations at different ages denoted on the $x-$axis. We use our preferred model with municipality and cohort fixed effects and family background controls (results are similar if we include municipality-specific trends). Consistent with results presented above, we find a large differential effect on hospitalizations before age one among younger siblings. However, this effect dissipates as we consider older ages. If anything, it appears that disease exposure in the first year of life is associated with a reduction in the number of hospitalizations at ages 3 to 4.

These results suggest that the idea that infections in early childhood are protective of later disease (e.g., **?**) appears to hold up our in our data. Moreover, our findings suggest that the differential adverse effects on long-term educational and economic outcomes of early-life disease exposure among younger siblings are *not* driven by worse health in later childhood. If anything, these children have slightly better health at older ages, but nevertheless experience penalties in terms of their long-run outcomes.

In Figure 5 examines differential effects on educational attainment and labor market outcomes measured at ages 22 to 30. The adverse effect on high school graduation emerges at ages 19 and 20 and stays relatively stable at older ages, while the effect on college graduation appears at age 23 and becomes larger in magnitude from age 24 onward. As for labor force participation and income, we observe a positive differential effect at ages 23 to 25, which is consistent with selection out of higher education and into work. The negative differential effect on these outcomes begins around age 26.

Heterogeneous effects across sub-groups. Panel (b) of Figure 4 presents results from models estimated on different sub-samples of the data for our main short-run outcome, the number of respiratory disease hospitalizations by age one. 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 (birth weight above and below the median, as well as low-birth-weight and non-low-birth-weight children), and the gender composition of the siblings. We find similar estimates across all sub-groups, with one important exception: low birth weight children (those whose birth weight is below 2,500 grams) experience a much larger differential increase in respiratory disease hospitalizations when they are the younger siblings.

We also explore heterogeneity in effects on long-term outcomes across these sub-groups in Figure 6. While the confidence intervals are mostly overlapping, it appears that the differential adverse effect on educational attainment is particularly pronounced among low-birth-weight children. Interestingly, the analysis of labor market outcomes reveals a positive differential effect of early life disease exposure among younger siblings when the younger siblings are girls while the older siblings are boys.

Sensitivity of results.

5 Conclusion

References

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

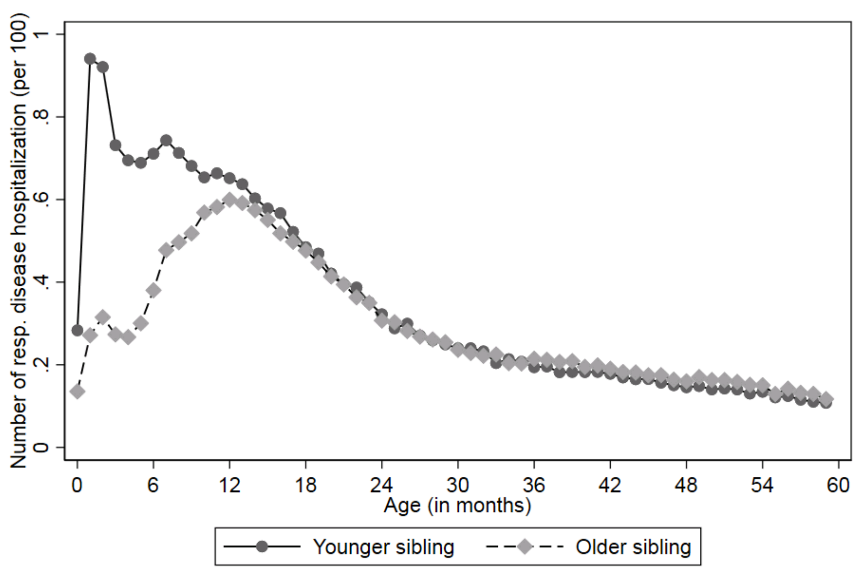
Currie, Janet and Douglas Almond, "Human capital development before age five," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1315–1486.

6 Figures

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

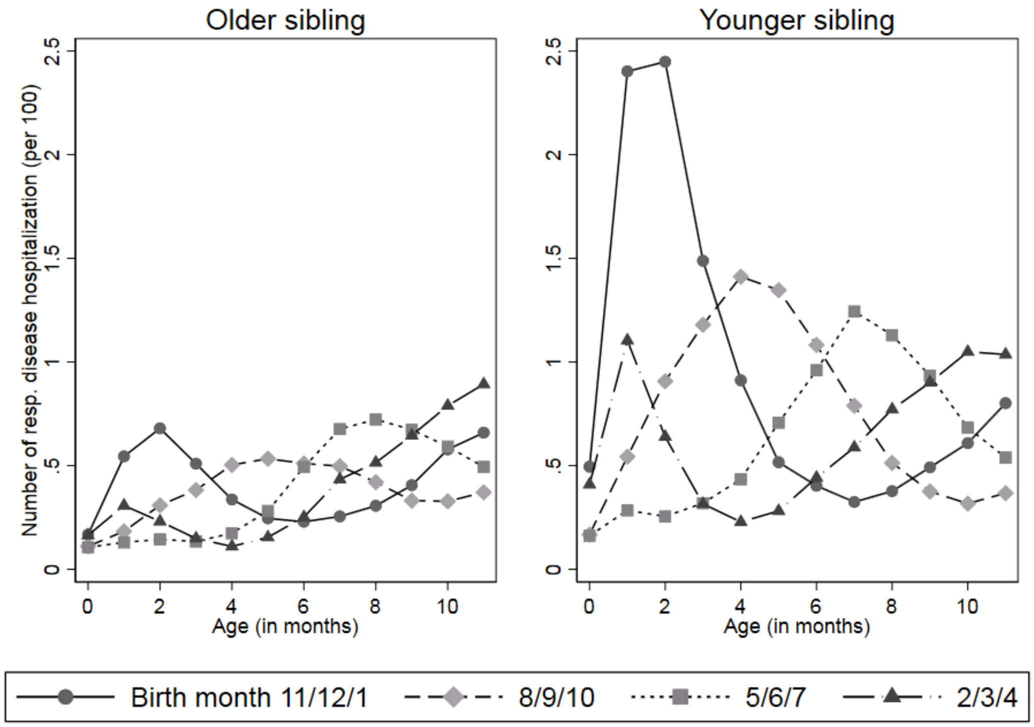
(a) During First Year of Life

(b) During First Five Years of Life



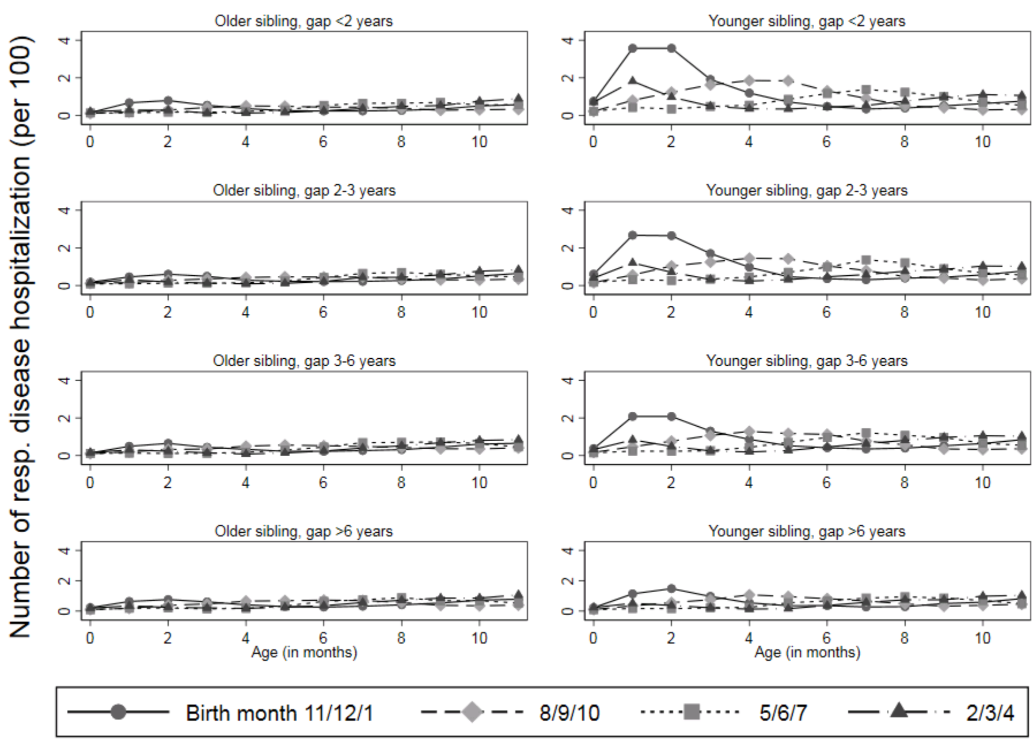
Notes: These figures plot the number of hospitalizations with respiratory disease diagnoses (per 100 children) by month of age, separately for older and younger siblings in our data.

Figure 2: Number of Respiratory Disease 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 disease 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.

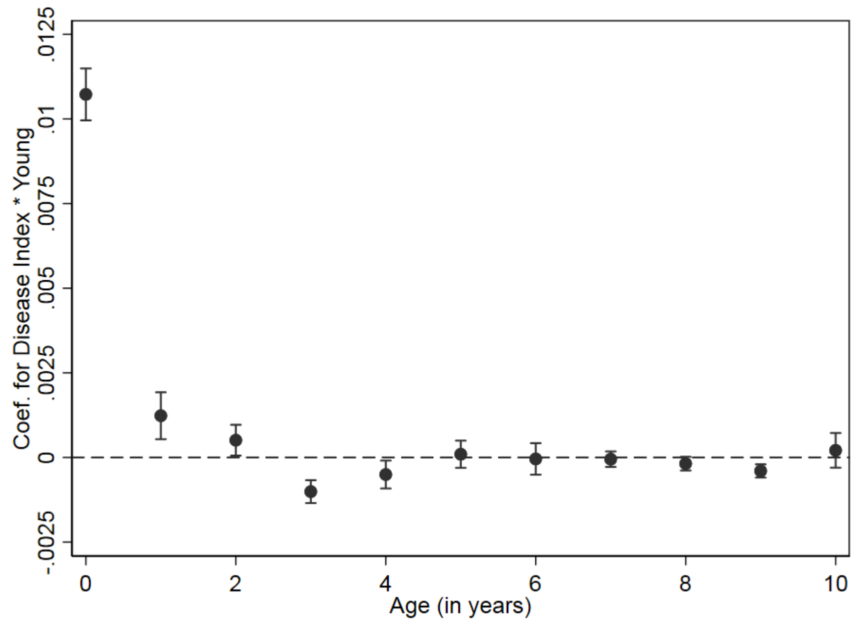
Figure 3: Number of Respiratory Disease Hospitalizations per 100 Children, by Child Age in Months, Season of Birth, and Birth Spacing, Older versus Younger Siblings



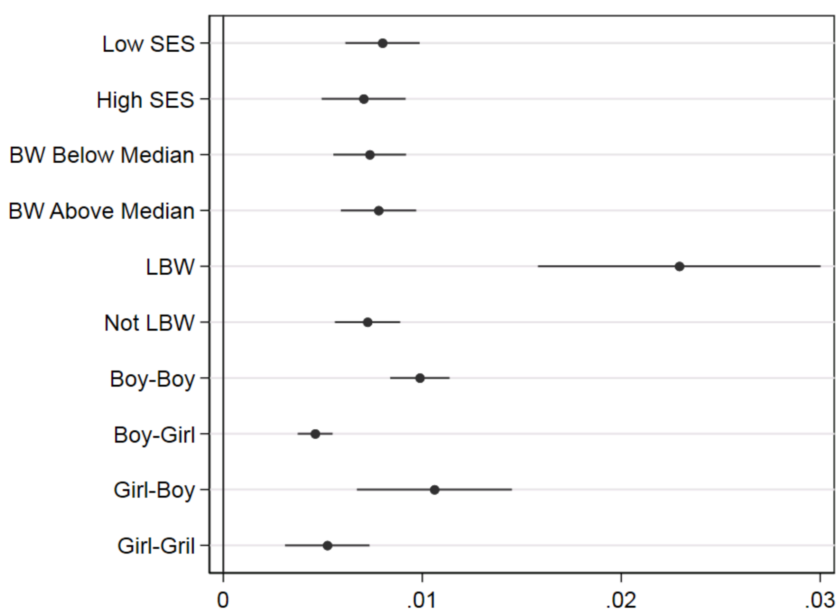
Notes: These figures plot the number of hospitalizations with respiratory disease diagnoses (per 100 children) by month of age and by the season of birth of the child, separately for older and younger siblings with different birth spacing gaps in our data.

Figure 4: Differential Effects of the Disease Exposure Index on the Annual Number of Younger Siblings' Respiratory Disease Hospitalizations, by Age of Observation and by Sub-Group

(a) By Age of Observation

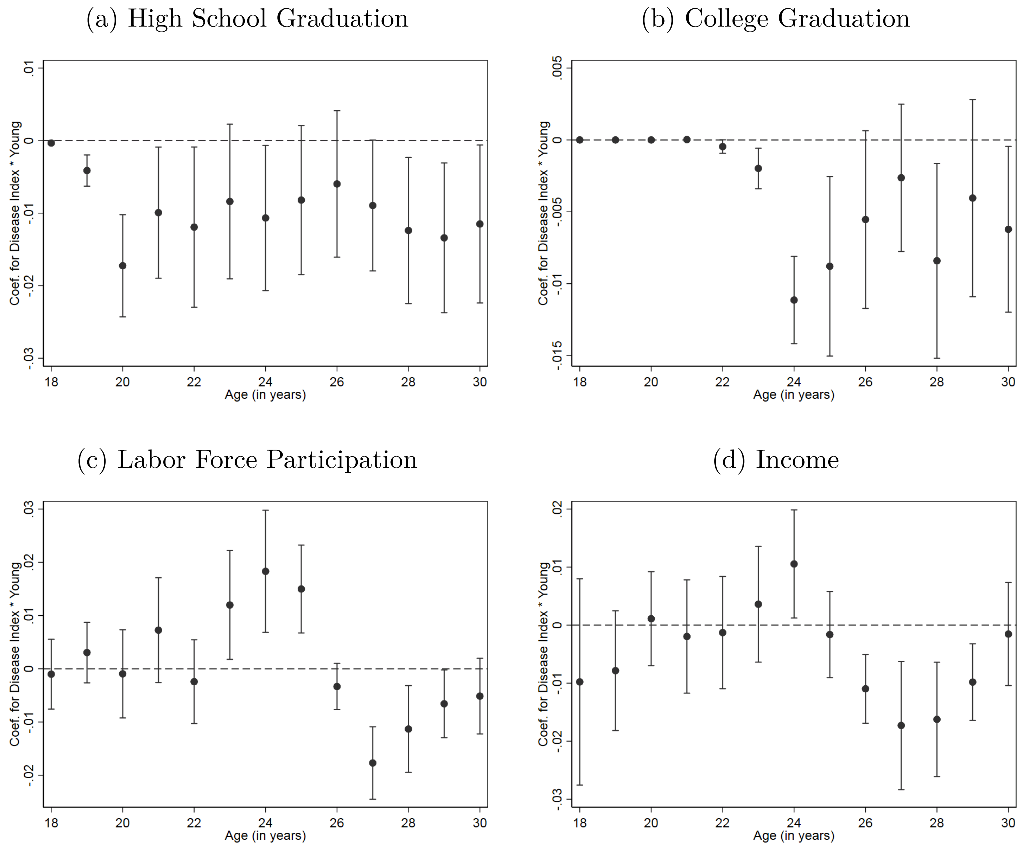


(b) By Sub-Group



Notes: Sub-figure (a) plots the coefficients and 95% confidence intervals on the interaction term between the disease index and the younger sibling indicator from model (1), using as outcomes the annual number of hospitalizations with respiratory disease diagnoses measured at ages specified on the x-axis. Sub-figure (b) plots these coefficients and 95% confidence intervals for this outcome measured in the first year of life, from models estimated on different sub-samples as specified on the y-axis.

Figure 5: Differential Effects of the Disease Exposure Index on Younger Siblings' Long-Run Outcomes, by Age of Observation



(a) High School Graduation

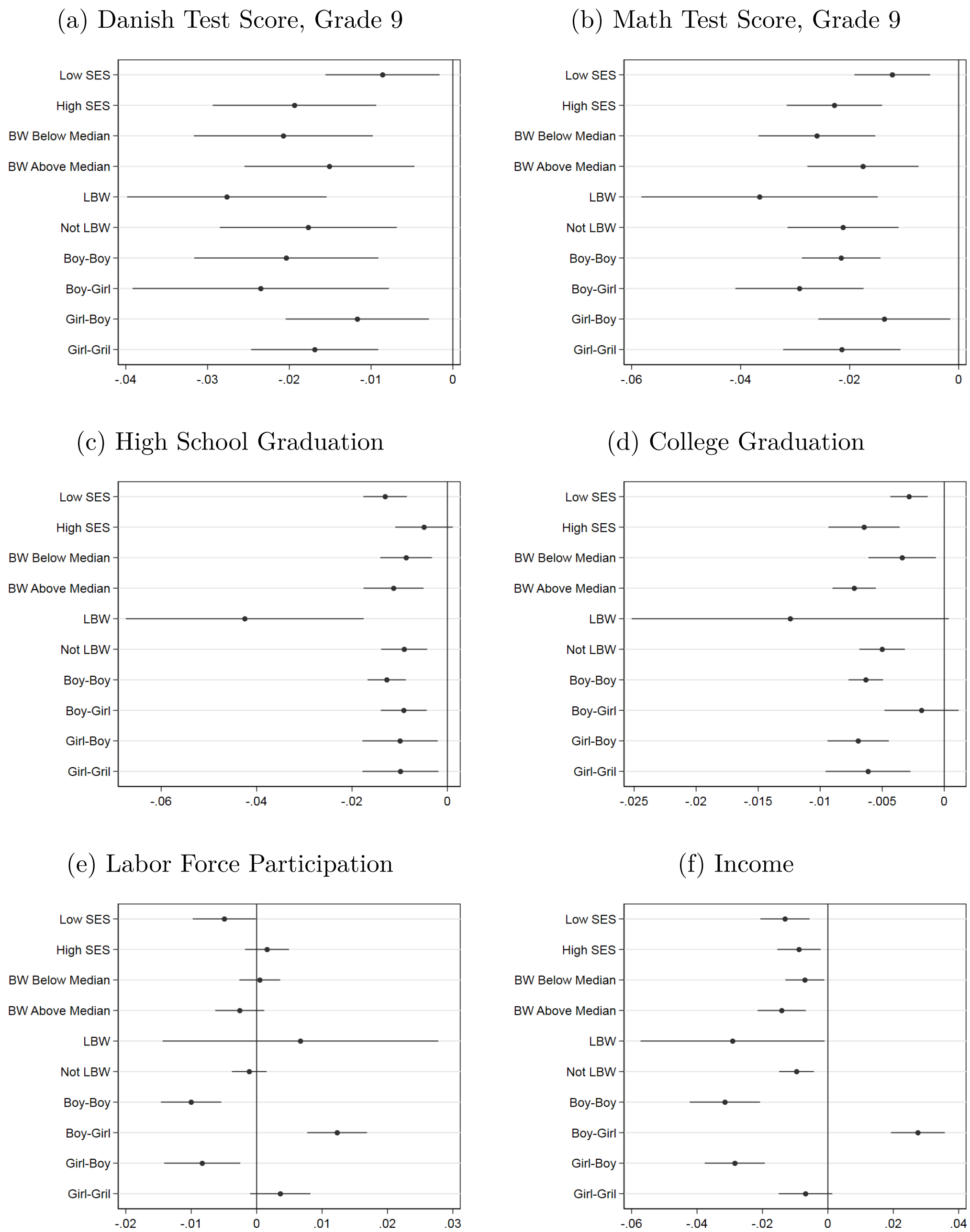
(b) College Graduation

(c) Labor Force Participation

(d) Income

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.

Figure 6: Heterogeneous Effects of the Disease Exposure Index on Younger Siblings' Outcomes



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), for different outcomes, across different subsamples as specified on the y-axes.

7 Tables

Table 1: Variable Means

	Older Siblings	Younger Siblings
Child Characteristics and Outcomes		
Male Child	0.514	0.514
Birth Weight (grams)	3418.707	3575.573
Respiratory Disease Exposure Index	2.775	2.866
Number of Respiratory Disease Hospitalizations by Age 1	0.047	0.089
Log Income, Age 26	10.496	10.473
In Labor Force, Age 26	0.610	0.602
High School Degree, Age 24	0.772	0.761
College Degree, Age 24	0.055	0.048
Danish Test Score, Grade 9	0.126	0.010
Math Test Score, Grade 9	0.140	0.017
Family Background Characteristics		
Mother's Age at Childbirth	26.631	30.177
Mother is Foreign-Born	0.071	0.071
Mother has High School Degree	0.739	0.771
Mother has College Degree	0.089	0.104
Parents are Married/Cohabiting	0.796	0.915
Log Household Income	11.408	11.581
Observations	667774	667774

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. Income variables are reported in 2010 $USD. 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 1986–2000. Maternal educational attainment and parental marital/cohabiting status are measured in the year of childbirth, while household income is measured in the year before childbirth.

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

	Respiratory Disease Hospitalizations in First Year of Life					
	(1)	(2)	(3)	(4)	(5)	(6)
Younger	0.043***		0.043***	0.006***	0.017***	0.016***
	(0.006)		(0.006)	(0.001)	(0.002)	(0.002)
Disease index		0.019***	0.019***	0.013***	0.011***	0.009***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Younger x disease index				0.012***	0.012***	0.012***
				(0.000)	(0.000)	(0.000)
Municipality FEs	Yes	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes	Yes
Family Background Controls	No	No	No	No	Yes	Yes
Municipality Trends	No	No	No	No	No	Yes
Observations	1,335,548	1,335,548	1,335,548	1,335,548	1,289,886	1,289,886
Mean	0.068	0.068	0.068	0.068	0.069	0.069
25th to 75th pctile effect size				0.022	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 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 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. All specifications include municipality, year of birth, and month of birth fixed effects. Columns (5)–(8) also include the following family background controls: indicator for child gender, 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. Columns (6) and (8) include municipality-specific linear trends. Columns (7) and (8) include mother fixed effects. 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 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)	(5)
Younger	-0.261***		-0.261***	-0.218***	-0.219***
	(0.012)		(0.012)	(0.019)	(0.023)
Disease index		-0.014***	-0.014***	-0.008	-0.009
		(0.004)	(0.004)	(0.006)	(0.007)
Younger x disease index				-0.014**	-0.013*
				(0.006)	(0.007)
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	No	No	No	Yes
Observations	532,473	532,473	532,473	532,473	532,473
Mean	0.104	0.104	0.104	0.104	0.104
25th to 75th pctile effect size				-0.024	-0.023

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 1986–2000. 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 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)	(5)
Younger	-0.272***		-0.272***	-0.229***	-0.227***
	(0.015)		(0.015)	(0.009)	(0.015)
Disease index		-0.008**	-0.007**	-0.001	0.005
		(0.003)	(0.003)	(0.004)	(0.005)
Younger x disease index				-0.014***	-0.014***
				(0.004)	(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	No	No	No	Yes
Observations	526,010	526,010	526,010	526,010	526,010
Mean	0.125	0.125	0.125	0.125	0.125
25th to 75th pctile effect size				-0.024	-0.025

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 1986–2000. 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 Disease Exposure Index in First Year of Life on Likelihood of High School Graduation, Younger versus Older Siblings

	Graduated High School by Age 24				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.083***		-0.083***	-0.070***	-0.064***
	(0.004)		(0.004)	(0.002)	(0.004)
Disease index		-0.001	-0.000	0.003*	0.007***
		(0.002)	(0.002)	(0.002)	(0.002)
Younger x disease index				-0.005***	-0.007***
				(0.001)	(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	No	No	No	Yes
Observations	341,740	341,740	341,740	341,740	341,740
Mean	0.770	0.770	0.770	0.770	0.770
25th to 75th pctile effect size				-0.007	-0.010

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is an indicator for graduating high school by age 24. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table 6: Effect of Disease Exposure Index in First Year of Life on Likelihood of College Graduation, Younger versus Older Siblings

	Graduated College by Age 24				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.028***		-0.028***	-0.017***	-0.014***
	(0.004)		(0.004)	(0.005)	(0.004)
Disease index		0.001	0.001	0.004**	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
Younger x disease index				-0.004**	-0.005***
				(0.002)	(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	No	No	No	Yes
Observations	341,740	341,740	341,740	341,740	341,740
Mean	0.050	0.050	0.050	0.050	0.050
25th to 75th pctile effect size				-0.006	-0.007

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is an indicator for graduating college by age 24. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table 7: Effect of Disease Exposure Index in First Year of Life on Labor Force Participation at Age 26, Younger versus Older Siblings

	Labor Force Participation at Age 26				
	(1)	(2)	(3)	(4)	(5)
Younger	0.023***		0.023***	0.026***	0.025***
	(0.002)		(0.002)	(0.006)	(0.005)
Disease index		0.004	0.004	0.005	0.005
		(0.004)	(0.004)	(0.004)	(0.004)
Younger x disease index				-0.001	-0.001
				(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	No	No	No	Yes
Observations	230,434	230,434	230,434	230,434	230,434
Mean	0.612	0.612	0.612	0.612	0.612
25th to 75th pctile effect size				-0.002	-0.002

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is an indicator for being in the labor force at age 26. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table 8: Effect of Disease Exposure Index in First Year of Life on Log Income at Age 26, Younger versus Older Siblings

	Log Income at Age 26				
	(1)	(2)	(3)	(4)	(5)
Younger	0.018**		0.018**	0.053***	0.046***
	(0.008)		(0.008)	(0.011)	(0.010)
Disease index		0.008*	0.008*	0.016***	0.019***
		(0.004)	(0.004)	(0.004)	(0.004)
Younger x disease index				-0.014***	-0.011***
				(0.003)	(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	No	No	No	Yes
Observations	228,003	228,003	228,003	228,003	228,003
Mean	10.502	10.502	10.502	10.502	10.502
25th to 75th pctile effect size				-0.018	-0.014

Notes: See notes under Table 2 for more details about the specifications and variables. The outcome is the natural log of gross income at age 26, converted into 2010 USD$. Standard errors are clustered on the child's municipality of birth. Significance levels: * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

A Appendix Tables

Table A1: Disease Exposure Index and *Predicted* Respiratory Disease Hospitalizations in First Year of Life, Younger versus Older Siblings

	Predicted Respiratory Disease Hospitalization in First Year of Life				
	(1)	(2)	(3)	(4)	(5)
	Pr_hosp_resp	Pr_hosp_resp	Pr_hosp_resp	Pr_hosp_resp	Pr_hosp_resp
Younger	0.008		0.008	-0.093**	-0.123**
	(0.041)		(0.041)	(0.044)	(0.053)
Disease index		0.032**	0.032**	0.016**	-0.022**
		(0.013)	(0.013)	(0.006)	(0.011)
Younger x disease index				0.032*	0.042**
				(0.017)	(0.019)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Municipality Trends	No	No	No	No	Yes
Observations	1,262,416	1,262,416	1,262,416	1,262,416	1,262,416
Mean	6.864	6.864	6.864	6.864	6.864
25th to 75th pctile effect size				0.059	0.077

Notes: Each column in the table presents results from estimating different versions of model (1). The outcome is the child's *predicted* number of respiratory disease hospitalizations, which is predicted after regressing the child's number of hospitalizations with a respiratory disease primary diagnosis in the first year of life on the following family background control variables: indicator for child gender, 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. 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 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 A2: 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	0.032***	0.016***	0.016***	0.013***	0.017***
	(0.007)	(0.002)	(0.004)	(0.003)	(0.003)
Disease index	0.006**				
	(0.002)				
Younger x disease index	0.014***				
	(0.001)				
Disease index (13-35 months)		0.005***			
		(0.001)			
Younger x disease index (13-35 months)		0.006***			
		(0.000)			
Disease index (36-71 months)		0.002***			
		(0.001)			
Younger x disease index (36-71 months)		0.005***			
		(0.001)			
Disease index (-12 to +24 months around old sibling)			-0.001***		
			(0.000)		
Younger x disease index (-12 to +24 months around old sibling)			0.017***		
			(0.001)		
Disease index (all diagnoses)				0.006***	
				(0.001)	
Younger x disease index (all diagnoses)				0.009***	
				(0.000)	
Disease index (children with any diagnoses)					0.012***
					(0.001)
Younger x disease index (children with any diagnoses)					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
Mother FEs	Yes	No	No	No	No
Observations	1,262,416	1,289,886	1,234,741	1,289,886	1,289,886
Mean	0.069	0.069	0.069	0.069	0.069
25th to 75th pctile effect size	0.026	0.005	0.026	0.023	0.021

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 during the first year of the child's life. Column (1) adds maternal fixed effects to our baseline model. Column (2) uses our baseline model but includes two respiratory disease exposure indices calculated using hospitalizations among children aged 13–35 months and 36–71 months separately. Column (3) uses a respiratory disease exposure index that is calculated using children whose age is between 12 months less to 24 months more than the child's older siblings. When calculating this index for the older siblings themselves, we assume that they have an older sibling with the same birth gap. We restrict the analysis sample to siblings with a birth gap of less than 96 months in column (3). Column (4) uses a disease index in which we include hospitalizations with primary and non-primary diagnoses for respiratory conditions. Column (5) uses a disease index in which we calculate the number of children with at least one respiratory disease diagnosis (i.e., counting children and not 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 A3: Robustness of Results on 9th Grade Danish Test Score

	9th Grade Danish Test Score				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.154***	-0.201***	-0.284***	-0.210***	-0.202***
	(0.010)	(0.026)	(0.014)	(0.022)	(0.024)
Disease index	-0.005				
	(0.006)				
Younger x disease index	-0.003				
	(0.003)				
Disease index (13-35 months)		-0.005			
		(0.003)			
Younger x disease index (13-35 months)		-0.004			
		(0.003)			
Disease index (36-71 months)		-0.002			
		(0.006)			
Younger x disease index (36-71 months)		-0.018			
		(0.011)			
Disease index (-12 to +24 months around old sibling)			0.013***		
			(0.001)		
Younger x disease index (-12 to +24 months around old sibling)			0.013***		
			(0.003)		
Disease index (all diagnoses)				-0.005**	
				(0.003)	
Younger x disease index (all diagnoses)				-0.010**	
				(0.005)	
Disease index (children with any diagnoses)					-0.011*
					(0.006)
Younger x disease index (children with any diagnoses)					-0.018**
					(0.008)
Municipality FEs	Yes	Yes	Yes	Yes	Yes
YoB+MoB FEs	Yes	Yes	Yes	Yes	Yes
Family Background Controls	Yes	Yes	Yes	Yes	Yes
Mother FEs	Yes	No	No	No	No
Observations	530,312	695,605	664,035	695,605	695,605
Mean	0.105	0.074	0.081	0.074	0.074
25th to 75th pctile effect size	-0.005	-0.017	0.019	-0.025	-0.030

Notes: See notes under Appendix Table A2 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 1986–2000. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A4: Robustness of Results on 9th Grade Math Test Score

	9th Grade Math Test Score				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.143***	-0.212***	-0.301***	-0.221***	-0.208***
	(0.011)	(0.020)	(0.016)	(0.018)	(0.018)
Disease index	0.003				
	(0.006)				
Younger x disease index	0.001				
	(0.003)				
Disease index (13-35 months)		-0.002			
		(0.002)			
Younger x disease index (13-35 months)		-0.007***			
		(0.002)			
Disease index (36-71 months)		-0.003			
		(0.005)			
Younger x disease index (36-71 months)		-0.008			
		(0.008)			
Disease index (-12 to +24 months around old sibling)			0.042***		
			(0.003)		
Younger x disease index (-12 to +24 months around old sibling)			0.017***		
			(0.002)		
Disease index (all diagnoses)				-0.001	
				(0.002)	
Younger x disease index (all diagnoses)				-0.009*	
				(0.005)	
Disease index (children with any diagnoses)					-0.005
					(0.004)
Younger x disease index (children with any diagnoses)					-0.018***
					(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
Mother FEs	Yes	No	No	No	No
Observations	523,878	690,426	659,216	690,426	690,426
Mean	0.127	0.085	0.097	0.085	0.085
25th to 75th pctile effect size	0.002	-0.008	0.024	-0.022	-0.030

Notes: See notes under Appendix Table A2 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 1986–2000. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A5: Robustness of Results on High School Graduation by Age 24

	Graduated High School by Age 24				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.040***	-0.065***	-0.087***	-0.063***	-0.063***
	(0.008)	(0.003)	(0.006)	(0.004)	(0.003)
Disease index	0.002				
	(0.004)				
Younger x disease index	0.002				
	(0.002)				
Disease index (13-35 months)		0.002			
		(0.001)			
Younger x disease index (13-35 months)		-0.003*			
		(0.002)			
Disease index (36-71 months)		-0.007			
		(0.005)			
Younger x disease index (36-71 months)		0.002			
		(0.005)			
Disease index (-12 to +24 months around old sibling)			0.003***		
			(0.000)		
Younger x disease index (-12 to +24 months around old sibling)			0.003***		
			(0.001)		
Disease index (all diagnoses)				-0.000	
				(0.001)	
Younger x disease index (all diagnoses)				-0.004***	
				(0.001)	
Disease index (children with any diagnoses)					0.001
					(0.002)
Younger x disease index (children with any diagnoses)					-0.006***
					(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
Mother FEs	Yes	No	No	No	No
Observations	323,984	434,528	414,032	434,528	434,528
Mean	0.772	0.770	0.775	0.770	0.770
25th to 75th pctile effect size	0.002	0.001	0.004	-0.008	-0.008

Notes: See notes under Appendix Table A2 for more details about the specifications and variables. The outcome is an indicator for graduating high school by age 24. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A6: Robustness of Results on College Graduation by Age 24

	Graduated College by Age 24				
	(1)	(2)	(3)	(4)	(5)
Younger	-0.011**	-0.016***	-0.024***	-0.017***	-0.016***
	(0.005)	(0.004)	(0.002)	(0.003)	(0.003)
Disease index	0.002				
	(0.002)				
Younger x disease index	-0.002				
	(0.002)				
Disease index (13-35 months)		0.001*			
		(0.001)			
Younger x disease index (13-35 months)		-0.002***			
		(0.001)			
Disease index (36-71 months)		0.000			
		(0.002)			
Younger x disease index (36-71 months)		-0.001			
		(0.002)			
Disease index (-12 to +24 months around old sibling)			0.004***		
			(0.001)		
Younger x disease index (-12 to +24 months around old sibling)			-0.002***		
			(0.001)		
Disease index (all diagnoses)				0.002	
				(0.001)	
Younger x disease index (all diagnoses)				-0.003***	
				(0.001)	
Disease index (children with any diagnoses)					0.003
					(0.002)
Younger x disease index (children with any diagnoses)					-0.004***
					(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
Mother FEs	Yes	No	No	No	No
Observations	323,984	434,528	414,032	434,528	434,528
Mean	0.051	0.053	0.054	0.053	0.053
25th to 75th pctile effect size	-0.003	-0.001	-0.003	-0.006	-0.006

Notes: See notes under Appendix Table A2 for more details about the specifications and variables. The outcome is an indicator for graduating college by age 24. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.

Table A7: Robustness of Results on Labor Force Participation at Age 26

	Labor Force Participation at Age 26				
	(1)	(2)	(3)	(4)	(5)
Younger	0.032**	0.029***	0.034***	0.027***	0.025***
	(0.016)	(0.007)	(0.005)	(0.006)	(0.006)
Disease index	0.017***				
	(0.005)				
Younger x disease index	-0.011***				
	(0.004)				
Disease index (13-35 months)		0.001			
		(0.001)			
Younger x disease index (13-35 months)		0.002			
		(0.001)			
Disease index (36-71 months)		0.009***			
		(0.003)			
Younger x disease index (36-71 months)		-0.010*			
		(0.006)			
Disease index (-12 to +24 months around old sibling)			-0.010***		
			(0.001)		
Younger x disease index (-12 to +24 months around old sibling)			-0.006***		
			(0.002)		
Disease index (all diagnoses)				0.003**	
				(0.002)	
Younger x disease index (all diagnoses)				-0.001	
				(0.001)	
Disease index (children with any diagnoses)					0.006***
					(0.002)
Younger x disease index (children with any diagnoses)					-0.001
					(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
Mother FEs	Yes	No	No	No	No
Observations	212,858	314,790	300,525	314,790	314,790
Mean	0.610	0.607	0.607	0.607	0.607
25th to 75th pctile effect size	-0.014	-0.008	-0.007	-0.003	-0.002

Notes: See notes under Appendix Table A2 for more details about the specifications and variables. The outcome is an indicator for being employed at age 26. 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 Log Income at Age 26

	Log Income at Age 26				
	(1)	(2)	(3)	(4)	(5)
Younger	0.021	0.050***	0.042***	0.044***	0.043***
	(0.023)	(0.014)	(0.009)	(0.011)	(0.012)
Disease index	0.029*				
	(0.017)				
Younger x disease index	-0.007				
	(0.005)				
Disease index (13-35 months)		0.002			
		(0.002)			
Younger x disease index (13-35 months)		0.001			
		(0.003)			
Disease index (36-71 months)		0.020***			
		(0.006)			
Younger x disease index (36-71 months)		-0.027**			
		(0.012)			
Disease index (-12 to +24 months around old sibling)			-0.012***		
			(0.001)		
Younger x disease index (-12 to +24 months around old sibling)			-0.013***		
			(0.003)		
Disease index (all diagnoses)				0.006**	
				(0.002)	
Younger x disease index (all diagnoses)				-0.007***	
				(0.002)	
Disease index (children with any diagnoses)					0.012***
					(0.003)
Younger x disease index (children with any diagnoses)					-0.010***
					(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
Mother FEs	Yes	No	No	No	No
Observations	210,688	313,158	298,959	313,158	313,158
Mean	10.499	10.488	10.488	10.488	10.488
25th to 75th pctile effect size	-0.009	-0.021	-0.015	-0.014	-0.013

Notes: See notes under Appendix Table A2 for more details about the specifications and variables. The outcome is the log of gross personal income at age 26. Standard errors are clustered on the child's municipality of birth. Significance levels: * p<0.1 ** p<0.05 *** p<0.01.