NATIONAL
CENTER *for* ANALYSIS *of* LONGITUDINAL DATA *in* EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

AIR

Sibling Spillovers

Sandra E. Black
Sanni Breining
David N. Figlio
Jonathan Guryan
Krzysztof Karbownik
Helena Skyt Nielsen
Jeffrey Roth
Marianne Simonsen

WORKING PAPER 192 • June 2018

Sibling Spillovers

Sandra E. Black
University of Texas and NBER

Sanni Breining
Aarhus University

David N. Figlio
Northwestern University and NBER

Jonathan Guryan
Northwestern University and NBER

Krzysztof Karbownik
Northwestern University

Helena Skyt Nielsen
Aarhus University

Jeffrey Roth
University of Florida

Marianne Simonsen
Aarhus University

Contents

Acknowledgements

We are grateful to the Florida Departments of Education and Health for providing the de-identified, matched data used in this analysis. Figlio and Roth acknowledge support from the National Science Foundation and the Institute for Education Sciences (CALDER grant), Figlio acknowledges support from the National Institute of Child Health and Human Development and the Bill and Melinda Gates Foundation, and Black acknowledges support from the Research Council of Norway through its Centres of Excellence Scheme, FAIR project No 262675. We appreciate helpful feedback from Scott Carrell, Dennis Epple, Richard Romano, and seminar and conference participants at AEFP, APPAM, CALDER, CESifo, BER Summer Institute Children's Program, SOLE/EALE and the University of Wisconsin. We would also like to thank Mikkel Aagaard Houmark for excellent research assistance. The conclusions expressed in this paper are those of the authors and do not represent the positions of the Florida Departments of Education and Health, Statistics Denmark or those of our funders.

This research was supported by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER), which is funded by a consortium of foundations. For more information about CALDER funders, see www.caldercenter.org/about-calder. All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders or the institutions to which the authors are affiliated.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any opinions, findings, and conclusions expressed in these papers are those of the authors and do not necessarily reflect the views of our funders.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street N.W., Washington, D.C. 20007
202-403-5796 • www.caldercenter.org

Sibling Spillovers

Sandra E. Black, Sanni Breining, David N. Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, Marianne Simonsen

CALDER Working Paper No. 192
June 2018

Abstract

It is notoriously difficult to identify peer effects within the family because of the common shocks and reflection problems. We make use of a novel identification strategy and unique data in order to gain some purchase on this problem. We employ data from the universe of children born in Florida between 1994 and 2002 and in Denmark between 1990 and 2001, which we match to school and medical records. To address the identification problem, we examine the effects of having a sibling with a disability. Utilizing three-plus-child families, we employ a differences-in- differences research design which makes use of the fact that birth order influences the amount of time that a child spends in early childhood with their siblings, disabled or not. We observe consistent evidence in both locations that the second child in a family is differentially affected when the third child is disabled.

1 Introduction

Brothers and sisters share a bond that is unlike any other relationship - they typically grow up in the same home, with the same parents and similar genetics, and experience life events together. It would not be surprising, then, to learn that siblings have important influences on children's outcomes. This influence might manifest in multiple ways. Siblings might teach each other directly; they might also model both good and bad behaviors. And since siblings also typically grow up in the same household raised by the same parents, they share limited parental resources. Time, attention, and money directed towards one child may be time, attention, and money diverted from another. Regardless of the mechanism by which they operate, the presence of spillovers across siblings implies that policies that affect one child in a family may have important effects on other children, suggesting a potential multiplier effect of policy's impacts.

That said, there is limited evidence regarding the causal role of siblings on children's outcomes.[1] Because siblings share many common traits and experiences, it is difficult to identify the effect of one sibling on another. In many ways, siblings are the ultimate peers, so estimating sibling spillovers faces all the same difficulties associated with estimating peer effects (e.g. Manski 1993; 2000; Moffitt 1999; Brock and Durlauf 2001; 2010; Epple and Romano 2010). Due to this, it is difficult to know how much of the cross-sibling correlations in economic outcomes (e.g., Solon 1999; Raaum et al. 2006; Mazumder 2008; Björklund and Jäntti 2012) are due to common shocks and unobservables versus sibling spillovers.

This paper makes use of population-level data from two contexts and a unique identification strategy to estimate the causal effect of a sibling health shock on educational outcomes of other children in the family. We use disability status as a vehicle through which we can study sibling spillovers; specifically, we examine the effect of having a sibling with a disability on test scores of older children, measured when they are in elementary and middle school. Children with disabilities require different kinds of investments of time, attention and financial resources from parents.[2] Any resulting diversion of these inputs may affect the development of other children in the family. At the same time, because siblings interact so frequently, disabled siblings may have a direct effect on children's outcomes. In addition to providing insight into sibling spillovers more generally, the case of sibling disability is itself economically significant, as millions of families are affected by disabilities of the type that we consider. In 2012-13, for instance, in the United States alone over 6.4 million

[1] As we discuss later, notable exceptions include Fletcher et al. (2012); Breining (2014); Nicoletti and Rabe (2014); Parman (2015); Yi et al. (2015); Adhvaryu and Nyshadham (2016); Breining et al. (2016b); Alsan (2017); Bingley et al. (2017); Ozier (2017); Qureshi (2017); Joensen and Nielsen (2018); and Qureshi (2018).

[2] An extensive literature documents the effects of child disability on parents. The increased complexity of parenting children with disabilities affects parental labor supply and the likelihood that families rely on public assistance (Salkever 1982; Crowe 1993; Kimmel 1998; Powers 2003; Noonan et al. 2005; Deshpande 2016b). Having a disabled child affects the likelihood that parents divorce or live apart (Reichman et al. 2004; Kvist et al. 2013), decreases social participation (Seltzer et al. 2001), and reduces maternal health (Burton et al. 2008) and grandparental well-being (Mitchell 2007). Our estimates will pick up both the direct effects in terms of parental time and resources as well as these indirect effects on children through their parents labor market, marital status, and health. In addition, a number of papers have documented the effects of childhood disability on one's own human capital accumulation (e.g. Currie and Stabile 2006; Fletcher and Wolfe 2008) and the role of welfare transfers for disabled youth (Deshpande 2016a). Stabile and Allin (2012) identify a number of economic costs of childhood disability.

children aged 3 to 21 (12.9 percent of all students) were supported under Part B of the Individuals with Disabilities Education Act (NCES 2016).

To address the difficult identification problem, we consider families with *three or more children*, where a health shock (disability) occurs in the case of the third child. Because these shocks are not randomly-distributed across families, we propose an identification strategy that looks *within* families. Our identification strategy relies on the idea that, within a family, the first- and second-born children face differential exposure to the affected third sibling. This differential exposure is related to the relative ordering of the two children; earlier-born children had more time in the family without the presence of the disabled third child, and are thus less "exposed." However, we cannot simply compare the first- and second-born children within these families; first-born and second-born children may be different for reasons unrelated to exposure to younger siblings.[3] By looking within families and then comparing these within-family differences to the same differences in families with a non-disabled third-born sibling, our approach allows us to separate the birth order effect from the effect of having a disabled sibling.

We are thus comparing the outcomes of the second-born (more affected) sibling to those of the first-born (less affected) sibling in families where the third-born sibling is disabled versus those where the third-born sibling is not. In doing so, we eliminate all biases associated with unobserved time-invariant family characteristics, and isolate the effects of siblings using a differences-in-differences, within-family comparison. While this approach allows us to estimate the effects of differential exposure to a disabled sibling, we are unable to estimate the total effect of this exposure since we cannot estimate a direct effect of the disabled sibling on the first-born child in the family.[4] Therefore, we think of our results as lower-bound estimates of the magnitude of the effect of having a disabled younger sibling.[5]

A key underlying assumption of our identification strategy is that the effect of birth order does not vary based on the presence of a disabled third sibling, other than through the differential exposure to the disabled third sibling. While we cannot test this directly, we do conduct a number of exercises to suggest that this might be a reasonable assumption. For one, we compare the differential effect of birth order on birth outcomes for children in families with and without a disabled third child; in other words, we can estimate the base difference-in-differences specification but for birth outcomes, which are predetermined relative to the birth of the third child, and therefore should not be affected by differential exposure to a disabled third child, rather than for outcomes at older ages, which occur after the third child's birth. As another way of testing whether there were differences in families with disabled and non-disabled third children prior to the birth of the third child, we

[3]There is a large literature on birth order effects, documenting differences in outcomes, with a particular focus on first-born children relative to others (e.g., Black et al. 2005; Conley and Glauber 2006; Price 2008; Booth and Kee 2009; Breining et al. 2016a).

[4]Any estimate of the general relationship between a disabled sibling and older siblings' outcomes will combine the true spillover effect with unobserved differences between families with and without disabled children.

[5]Two recent papers - Fletcher et al. (2012), using PSID-CDS, and Breining (2014), using Danish registry data – use cousin comparisons to study the correlations between childhood disabilities and sibling outcomes. Cousin comparisons reduce the omitted variables bias associated with differences across extended families, but within-extended-family differences may still remain.

investigate the special set of cases in which the third child was born after the second child already started school. We therefore examine whether there is any differential in the relative likelihood that first-born and second-born children have negative outcomes (including a disability diagnosis) by age five in families with disabled and non-disabled third children in that set of families. We find that first- and second-born siblings are equally likely to be considered disabled by age five in both types of families.

Because our research strategy is based on differential exposure to disabled younger siblings in early life, it is particularly important to observe disabilities that are noticeable to families early. We make use of administrative data that reveal when children first obtain services for their disabilities, and we concentrate on disabilities that are first identified in or before kindergarten. It is harder to separate the spillover effect from the common-shock effect for disabilities first diagnosed in the elementary school years because those disabilities are more likely to reflect unobserved family factors rather than a singular health or physical shock. Furthermore, as an alternative to measuring the effects of disabilities that appear in early childhood, we also consider the spillover effects of exposure to a third-born sibling with health problems at birth, such as congenital anomalies, abnormal conditions, low birth weight, prematurity, or very poor perinatal health. We find similar results regardless of whether we study the spillover effects of the third-born child's disability as recorded in early childhood or of the child's poor birth conditions.

We estimate these relationships using data from two data sources, drawn from two distinct locations: Florida and Denmark. In both cases, we make use of matched administrative datasets where we merge full birth cohorts with schooling data. The large sample sizes allow us to generate good statistical power while focusing on families with at least three children, a subset of which has a third child with disabilities. The data also allow us to follow children for whom we know birth outcomes and family structure into early adolescence to measure cognitive development.

The Florida data include birth records for every birth in the state between 1994 and 2002, merged with schooling records for all children who attended public schools within the state. The key outcome we measure in Florida is the state's criterion-referenced end-of-year standardized test, the Florida Comprehensive Assessment Test (FCAT), administered to children in grades 3-8 over the relevant time period (for more detail on these data, see Figlio et al. 2013 and Figlio et al. 2014).[6]

The Danish data include birth records for every birth in the country between 1990 and 2001, merged with medical patient register and school records. Our main outcome of interest in Denmark is the cohort-standardized grade from the ninth grade exit exam (GPA). These exam results are based on assessments by the student's teacher and by an external reviewer. While these exam scores do not capture exactly the same features as the Florida exams, they are widely interpreted to represent similar types of outcomes in the extant literature.

Including data from two very different contexts provides a number of advantages. For one, the United States and Denmark have very different educational systems, health care systems, and social welfare systems. To the degree to which we find comparable findings in these two quite different

[6]We pooled a child's FCAT observations across grades to reduce the potential consequences of measurement error. That said, our results are stable if we focus on a certain grade as opposed to pooling across multiple grades.

environments, this further increases our confidence that the results are externally valid and more generally reflect spillovers within the family, as opposed to reflecting something that is very context-specific. Second, disability is measured differently in the two contexts – on the education record in Florida and on the medical record in Denmark. This fact also aids in establishing the degree to which our findings are externally valid. Furthermore, having data from both settings provides greater opportunities to explore the mechanisms through which these findings might be operating because the two countries have quite different population characteristics and different data strengths. Thus, between the two sites, we are able to paint a more complete picture about the nature of the sibling spillovers that we seek to investigate.

In both locations, we find evidence consistent with there being a sibling spillover. Specifically, we observe that, in both Denmark and Florida, the second-born child in a family has worse outcomes (test scores in Florida, grade point average in Denmark) than does his or her older sibling when the third-born sibling is disabled, relative to the case in which the third-born sibling is not disabled. The magnitude of these differences is policy relevant; for example, in Florida it is about half of the observed relationship between an extra year of maternal education and children's test scores. These results are concentrated in cases in which the third child's disability is observed early – and therefore, presumably, more likely to affect older siblings in early childhood. Furthermore, the results are more pronounced in the case of physical disabilities – which are likely to be more visible early – rather than cognitive or behavioral disabilities – which are less likely to be manifested and diagnosed early.

As mentioned above, there are a number of potential mechanisms through which these sibling spillovers might occur. For one, siblings might have a direct effect on one another. But they also might have an indirect effect on one another through their effect on the allocation of parental resources – both of time and of money. Families with a child who requires additional attention, either in terms of time or in terms of costly services, have fewer resources to share with their other children. Analyses of the mechanisms by which sibling spillovers operate suggest that in addition to any direct effect siblings have on each other, the presence of a disabled child in a family appears to affect his or her siblings in part by changing parental allocations of time and financial resources away from the non-disabled children.

This study relates to a number of literatures, but most directly to work that examines how siblings indirectly or directly affect each other's health and educational outcomes. Though the empirical challenges associated with measuring causal effects among siblings has limited the amount of evidence on this topic, a number of recent studies have begun to build a body of evidence on how siblings affect each other within families. Some studies have generated evidence of sibling effects in developing countries (e.g. Yi et al. 2015; Adhvaryu and Nyshadham 2016; Alsan 2017; Ozier 2017; Qureshi 2018) or within the United States a century ago (Parman 2015).[7]

[7] Adhvaryu and Nyshadham (2016) evaluate an iodine supplementation program in Tanzania and find that there are effects on the treated children's siblings as well as the target child. Alsan (2017) examines an immunization campaign in Turkey and finds positive effects on the human capital of children not age-eligible for the program but whose siblings were eligible. Ozier (2017) evaluates a deworming intervention in Kenya and finds positive test score

More directly related to our work, other studies of developed countries have examined the effects of educational investments by older children on their younger siblings. These studies have tended to focus on changes in educational investments or educational choices that occurred during pre-adolescence or adolescence, and measure the effect on educational outcomes of younger siblings. The timing of the shock to the older sibling in these studies suggests it is plausible they measure the direct influence that older siblings' behavior has on younger siblings' behavior, in contrast to shocks to siblings earlier in life, as in our own study, which may be more likely to generate spillovers that operate through parents.[8]

Our work is one of the few papers to examine the effect of a health shock of a sibling during early childhood on the educational outcomes of other siblings in developed countries; the focus on health shocks early in life allows us to measure sibling spillovers at a time when parents likely have the greatest role in mediating the effects children have on each other within the household. The closest work to the current study is Breining (2014), which uses ordinary least squares and cousin fixed effects models and finds that having a younger sibling with attention deficit hyperactivity disorder (ADHD) significantly reduces first-born children's achievement test scores during high school, and Breining et al. (2016b), which uses a regression-discontinuity design to show that the siblings of children who were born just below the threshold to be deemed very low birthweight (VLBW) – and who received additional medical treatment very early in life – have higher language and math test scores in 9th grade on the order of 0.255 to 0.386 standard deviations, respectively. Similar results to Breining (2014) are reported by Fletcher et al. (2012) using PSID-CDS data and a broader set of health conditions. Our paper builds on this work, and is the first to examine the effects of sibling disability status on the educational success of children using within family design. We are also able to measure these effects in two different developed countries (with very different institutional structures), enabling us to better understand the role of institutions in mitigating these effects.

effects on siblings of children who were treated through the program. Qureshi (2018) uses gender segregation of schools in Pakistan to measure the effect of older siblings' schooling on their younger siblings, finding that increases in older sister's schooling has a positive effect on their younger brothers' literacy and schooling. Yi et al. (2015) examine twins in China and find that in twin pairs where one twin experienced a negative health shock at ages 0-3, parents made more health-related investments in the twin that experienced the early health shock and more education-related investments in the other twin, showing that parents can be a mechanism by which one sibling's health affect another's human capital. Parman (2015) studies the 1918 influenza pandemic in the United States and finds that families with a child born during the pandemic invested more in the education of their older children.

[8] Joensen and Nielsen (2018) study siblings in Denmark to estimate effects of an older sibling taking an advanced math-science course during high school on their younger siblings. They find that it increases the likelihood their younger sibling will take an advanced math-science course when they reach high school by 2-3 percentage points. Using data from North Carolina, Qureshi (2017) measures the effect of having an experienced classroom teacher on student achievement. She finds that having an experienced teacher has a positive effect on a child's younger siblings, but not on their older siblings. Qureshi interprets these findings as evidence that the spillover she measures operates through a direct effect of one sibling on another rather than through parents. Nicoletti and Rabe (2014) use a fixed-effects design to estimate the effect of improved academic achievement by an older sibling on the academic achievement of their younger sibling among children in England. They find that an increase in test scores at ages 16 has a modest positive effect on the test scores of younger siblings. In a context outside of education, Bingley et al. (2017) study conscription in Denmark and find that the random assignment of an older brother to serve in the military increases younger brothers' probability of serving in the military by 8 percent.

2 Data and descriptive statistics

The Florida data in our analysis are based on all singletons born in Florida between 1994 and 2002 and subsequently educated in Florida public schools. For the purposes of this research, Florida's education and health agencies matched children along three dimensions: first and last names, date of birth, and social security number. Figlio et al. (2013) describe details of the data and match quality; the match rate between birth and school records is nearly identical to that which would have been predicted using data from the American Community Survey. We observe sibling matches for the majority of the Florida data; Figlio et al. (2014) present the attributes of the sibling-matched data.

For Denmark, the key data source is the Danish Birth Register, which includes information on all individuals born in the period 1960-2010 from which we select a subsample of 1990 to 2001 births. For each child, the data include information on exact date of birth and various birth outcomes.[9] A unique identification number enables us to link children to their parents and siblings. Given this structure of the data and access to the date of birth, we can measure each individual's birth order and match the birth records to rich data from various administrative registers. This provides us with demographic characteristics of the parents such as age, educational attainment, labor market status, earnings, leave taking, and immigrant status.

We measure disability differently in the two settings. In Florida, disabilities are recorded on the school records, and are mutually exclusive categories.[10] The most common disabilities first observed by age five in Florida are speech impairment (48%), developmental delay (21%) and language impairment (17%). The most common disabilities first observed between ages five and ten in Florida are specific learning disabilities (46%), speech impairment (29%) and language impairment (9%). In Denmark, on the other hand, we observe disabilities recorded in the medical registries, and they are not mutually exclusive categories. The most common disabilities, among all children with any diagnoses, observed by age five in Denmark are congenital malformations and deformations of the musculoskeletal system (20%), congenital malformations of the circulatory system (10%), and congenital malformations of genital organs (9%); and the most common disabilities first observed between ages five and ten are behavioral and emotional disorders (29%), disorders of psychological development (14%) and congenital malformations of eye, ear, face and neck (8%).[11] As noted earlier, the fact that we observe different dimensions of disability in the different settings increases the potential for external validity of this study.

Our identification strategy requires us to observe the first several births in a family. Furthermore, because we want to have conditions as similar as possible between the first two births, we restrict the sample to the set of families for whom the father is the same in both cases. Since we make use of data from a longer range of birth years for Denmark than for Florida, we observe information

[9] Birth outcomes are only observed from 1994 forward.

[10] A child may have multiple disabilities and we observe all of these but there is always only one primary disability which we use in the analysis.

[11] Asthma is common among children, but we do not include this as a disability. Our results are robust to treating asthma as a disability.

on a larger number of families in Denmark than in Florida, despite the fact that Florida is more populous than Denmark.[12] There are 80,879 Florida families for whom we observe the first two children, and of these, we observe three or more children in 9,987 families. In Denmark, we observe the first two children in 134,277 families and the first three children in 28,581 families. Of the three-child families, 2,483 Florida families and 1,882 Denmark families have a third child observed with a disability by age five.

Table 1 presents descriptive statistics for our analysis population.[13] The table is divided into two panels, the top showing descriptive statistics for the Florida sample and the bottom for the Denmark sample. Column 1 shows descriptive statistics for all families with at least two children and the same father for the first two children, while column 2 shows descriptives for all families with at least three children and the same father for the first two children, a sample restriction we make to carry out our research design. We do not impose any restrictions on the disability status of the first two children in a family.

The Florida sample is demographically diverse. Twelve percent of the mothers in the Florida sample are African-American, and 23 percent are Hispanic.[14] Eight percent of sampled Florida mothers are high school dropouts, and 33 percent are college graduates. Note that the requirement that the first two children have the same father on the birth certificate makes this a more advantaged population than the overall population of Florida births.[15] The Danish sample is less racially and ethnically diverse than the Florida population, though 11 percent of the mothers in Denmark are immigrants. Denmark also has more high school dropout (22 percent) and fewer college graduate (24 percent) mothers in the analysis sample. Just as in Florida, the Danish estimation sample is comparatively privileged in terms of socio-economic background.[16] The age at first birth, however, is quite similar on average: 25.8 in Florida versus 26.3 in Denmark.

The sample of families with at least three children is somewhat different from the sample of families with at least two children (conditional on the first two having the same father in both samples). In Florida, families with at least three children are more likely to have African-American mothers (17 percent versus 12 percent), but have similar shares of Hispanic mothers. In Denmark, families with three or more children are more likely to have immigrant mothers (16 percent versus 11 percent). In Florida, mothers with less education are more likely to have families with three or more children (11 percent versus 8 percent high school dropout mothers, 29 percent versus 33 percent college graduate mothers), though the pattern is reversed in Denmark (26 percent versus 24 percent college graduate mothers).

[12] In Florida, we are restricted to only those families where we can observe the first three births during the nine-year birth records window, and we are further limited by the fact that siblings are only matched in a subset – albeit a large majority – of Florida counties. The United States also has a lower rate than Denmark of cases in which the father is the same for the first two births.

[13] The numbers of observations in our analyses sometimes differ from the numbers reported in this table, due to missing data on particular outcome variables for some individuals.

[14] African-American and Hispanic are not mutually-exclusive categories.

[15] Overall, 23 percent of all Florida births are to African-American mothers, 23 percent are to Hispanic mothers, 21 percent are to high school dropout mothers, and 20 percent are to college graduate mothers.

[16] Overall, 12 percent of all Danish births are to immigrant mothers; 26 percent are to high school dropout mothers, and 21 percent are to college graduate mothers.

Partly because all Florida births in the data come from a nine-year window, the average spacing between first and third births in Florida (4.6 years) is shorter than in Denmark (5.9 years). Birth spacing between the first and second born is also somewhat wider in Denmark than Florida, and in both places families that go on to have a third child are also likely to have first and second born children who are born closer together. On average, among those with sibling sets observed in the data, second born children arrive when the first born is 2.7 years old in Florida and 2.9 years old in Denmark. Among families with at least three children, that birth spacing is 2.2 years in Florida and 2.4 years in Denmark.

Columns 3 and 4 of Table 1 compare families with a third child who is disabled to families with a third child who is not disabled. In Florida, mothers with a disabled third child are less likely be African-American or Hispanic, though in Denmark, there is no difference in the share of mothers who are immigrants across these groups. In both Florida and Denmark, mothers with a disabled third child are less educated on average than mothers with a non-disabled third child. Mothers with a third disabled child are a bit younger when they give birth to their first child (23.5 versus 24.0 in Florida, 25.5 versus 25.7 in Denmark), and are less likely to be married at the time of birth, at least in Florida (70 percent versus 73 percent). These differences highlight that cross-family comparisons between children who have disabled siblings and those who do not may be confounded by differences in both observables and unobservables. This potential for confoundedness is what motivates our research design that relies on comparison within families and that controls for differences across families.

3 Differences-in-differences empirical strategy

Our research approach is to carry out a simple differences-in-differences design: We compare second versus first-born children, in families with disabled and non-disabled third children. The basic idea underlying the comparison is that in families with a disabled third-born child, second-born children spend a larger share of their early childhood exposed to a disabled sibling than first-borns do. However, first and second-born children may have different outcomes because of the direct effects of birth order. We therefore subtract off the first- versus second-born difference measured among families that have non-disabled third-born children as a way of separating the birth-order effect from the effect of differences in exposure to the third-born sibling.

Put differently, the comparisons in Table 1 make it clear that families with disabled children are different in some ways from families with non-disabled children. This means that comparisons between children who have disabled siblings and children who do not have disabled siblings may be confounded by any differences in unobservables. Our research design accounts for those differences across families by making comparisons within families, across first- and second-born siblings.

Our main regression equation is the following:

$$Y_{if} = \alpha_f + \beta_1 M_{if} + \beta_2 D_f M_{if} + \beta_3 X_{if} + \varepsilon_{if} \quad i = 1, 2 \tag{1}$$

where Y denotes a student's standardized test scores or grade point average (normalized to have mean zero and standard deviation one in population), D is an indicator variable taking the value one if the third child is observed disabled and zero otherwise, and M is an indicator variable taking the value one if the individual is second born, henceforth "more exposed", and zero if the individual is first born, or "less exposed". The indicator variable for third child disability (D) is not identified in this fixed effects model because it is constant across the first two births. The vector of covariates, X, includes indicator variables for year and month of birth as well as child gender. The index i indexes individuals and f indexes family, and α_f denotes the family fixed effect. Our main parameter of interest is β_2, which represents the difference in achievement gaps for more exposed versus less exposed siblings in families with and without a disabled third child, in a model in which family fixed effects net out time invariant differences between families with and without a disabled third child.

The key identifying assumption in a difference-in-differences approach is that there can be no differential trends between the treatment and control group in the absence of treatment. One concern is that, even before the third child is born, families with disabled third children may be trending in their risk of adverse outcomes in a different manner than are those without disabled third children. If children in families with disabled third children were already becoming progressively comparatively less successful for reasons that are unrelated to the third child's disability, this could bias our results and lead us to overestimate the effect of sibling spillovers.[17] We investigate this common trends assumption by studying birth weight and five-minute APGAR scores as outcomes of the analyses. While the identifying assumption cannot be tested directly, we would interpret potential "effects" on prior birth outcomes as an indication of confoundedness.

We estimate the main regression equation for the entire sample as well as for several subgroups defined by e.g. education of mother or type of disability. Furthermore, as another robustness check, we replace M by the number of years of exposure, which means that we identify the effect of interest off variation in birth spacing instead of birth order. In Section 4.1 we further explore an alternative treatment variable based on health conditions observed at the time of birth, rather than measured by disability status, in order to exclude the possibility that our results are driven by endogenously-determined disability identification during childhood.

4 Results: Exposure to a disabled younger sibling

The simple difference-in-difference results for our Florida sample are shown in Panel A of Table 2. Columns 1 and 2 show mean test scores for second- and first-born children, respectively, and column 3 shows the difference. The first and second rows show mean test scores for families with and without a disabled third-born child, respectively, and the third row shows the difference. We define disability

[17] Another, likely more minor concern, would be selective abortion where parents terminate pregnancy in anticipation of a disabled child. Although we are unable to address this issue empirically, due to data limitations, we believe that if substantial such effects should show up in family size. However, we do not observe that third child disability is correlated with either subsequent fertility or family size (completed family size in the case of Denmark).

based on the categorization in school records at kindergarten entry, around age five. As can be seen in the first and second column, first-born children score higher on standardized tests in third through eighth grade than do second-born children; this could be due to many reasons, such as differential exposure to undivided parental attention or longer exposure to parental influence, highlighted in the extensive birth order literature cited in the introduction. In families with disabled third-borns, the difference in test scores between first- and second-borns is 0.112 standard deviations, and in families with non-disabled third-borns, the difference is 0.064 standard deviations. The difference between those two differences, 0.048 standard deviations (p=0.036), is our estimate of the effect of additional exposure to a disabled younger sibling on third through eighth grade test scores. Standard errors are adjusted for clustering at the family level.

It is also interesting to note that children in families with disabled third-borns score lower than their counterparts in families with non-disabled third-borns. This pattern is apparent among both first and second-borns. First-borns in families with a disabled third-born sibling have test scores that are 0.040 standard deviations lower than first-borns in families with non-disabled third-born siblings. And second-borns in families with a disabled third-born sibling have test scores that are 0.089 standard deviations lower than second-borns in families with a non-disabled third-born sibling. The difference in those differences is also our estimate of the effect of additional exposure to a disabled younger sibling. It indicates that the gap between children from families with disabled and non-disabled younger siblings is greater among second-borns than among first-borns. This pattern is consistent with a causal effect of the additional exposure to a disabled child that second-borns experience relative to their older first-born siblings.

Panel B shows the parallel results for Denmark. Similar to the patterns shown for Florida, Danish first-borns score higher than second-borns, and Danish children in families with disabled siblings score lower than those with non-disabled siblings. The gap in grade point average between first-born and second-born children with disabled third-born siblings is 0.123 standard deviations. The gap in families with non-disabled third-born siblings is 0.089 standard deviations. The difference-in-differences estimate in Denmark is 0.034 standard deviations, somewhat smaller than the estimate in Florida, but still statistically significant (p=0.073). Smaller effects for Denmark may reflect the fact that the educational, health care, and social welfare systems are focused on alleviating the impact of disadvantages such as disabilities on the individual and the family.

Table 3 presents the same basic results but in a regression model with family fixed effects and controls for child gender and year and month of birth. Column 1 in row 1 of Table 3 presents the estimated effects of a third-born child identified as disabled by age five on a second-born sibling's outcomes relative to the first-born sibling's outcomes. These results are extremely similar to those presented in Table 2. The difference in differences model is appealing because it transparently shows the variation that identifies the effect of exposure to a disabled sibling. By comparing first-borns to second-borns, however, it ignores some information that might be useful for measuring the sibling spillover. Namely, there is variation across families in the spacing between first and second-born children. In families where the first and second-born children are spaced farther apart, the

difference in exposure to the third-born sibling is greater. To leverage this additional variation, we also estimate a model that includes family fixed effects, and an interaction between the difference in exposure to the third-child and an indicator for whether the family has a disabled third-born child. These results are presented in Columns 3 and 4 in Table 3.

In the subsequent two rows, we then augment the model to make the same comparisons in the case in which the third child is identified as disabled by age five versus the case in which the third child is identified as disabled at some point between ages five and ten. The purpose of this second specification is to gauge the degree to which relatively late-observed disabilities appear to differentially affect second-born versus first-born siblings. The answer appears to be that there is little evidence of a differential effect of late-observed disabilities. While the point estimates associated with early-observed disabilities are nearly identical regardless of whether we include or exclude an interaction between birth order and whether the third child was observed disabled between ages five and ten, the estimated differential for the late-observed group is a small fraction of that seen with the early-observed group, and far from statistically significant at conventional levels. For completeness, in the final row we also present results for cases in which the third-born child was observed disabled by age ten; this is simply a combination of those first observed disabled by age five and those first observed disabled between ages five and ten. As a consequence, it appears that the disabilities that have the biggest spillover, at least as identified using our particular strategy, are those that are observed early. This may be because these disabilities tend to be more severe, or at least more noticeable in very early childhood, and therefore more likely to have affected older siblings while they were comparatively young (Heckman 2006; Cunha and Heckman 2007).[18] Panel B shows the parallel results for the Denmark sample, which albeit smaller are similar to patterns documented for Florida with the exception of lower precision in the estimates utilizing spacing between siblings.

In addition, we have examined the effects of particular groups of disabilities (See Appendix Table A4); namely: physical or cognitive and behavioral. We find the largest deleterious effects in the case of physical disabilities. We also examine whether or not there are differences for reading versus mathematics; when we do so (Appendix Table A5), we find larger estimates for reading in comparison to mathematics in Florida. This is consistent with the possibility that home production contributes more to reading than to mathematics performance. On the other hand, in Denmark, the results are larger for mathematics than for reading, though neither is statistically significant. The point estimates in mathematics are very similar across the two countries.

4.1 Testing for unconfoundedness

We conduct a number of robustness and specification checks to verify that our identification strategy is valid. First, it may be the case that families with disabled third children were trending in a manner that is different from those without disabled third children even prior to the birth of the third child.

[18] In the Danish data we can directly address disability severity by investigating hospitalization records. Those observed disabled by age five have twice the rate of inpatient admission, three times the rate of outpatient admission, and slightly more emergency room admissions as do those first observed disabled between ages five and ten.

If children in families with disabled third children are becoming progressively comparatively less healthy or academically successful for reasons that are unrelated to the third child's disability, this could lead us to overestimate the effects of sibling spillovers.

While it is impossible to completely rule out this possibility, we can investigate whether we see these patterns when we look at birth outcomes of the first and second born children. Specifically, we consider four different birth outcomes observed in both Denmark and Florida: Birth weight in grams, log of birth weight, low birth weight (defined as birth weight less than 2500 grams), and five-minute APGAR scores (observed on a ten-point scale). Table 4 presents an analogous analysis to Table 3 – the coefficients on the interaction between the third sibling's disability status by age five and second-born status in the family in a model with family fixed effects and indicators for gender, year and month of birth.[19] Table 4 has fewer observations in Denmark than previous tables because we only observe birth outcomes for children born from 1994 forward. The results from Table 3 are similar when we restrict the Danish observations to be the same as those for which we can test for unconfoundedness in birth outcomes.

Importantly, we find no apparent relationship between third-born disability and the differential birth outcomes of second versus first-born siblings. Consider, for example, the estimated birth weight relationship. While we find that birth weight is slightly smaller for second-born versus first-born siblings in the case of a disabled (observed by age five) third-born sibling, 17 grams in Florida or 11 grams in Denmark, these differences are tiny relative to the estimated effects of a disabled third-born sibling on differential test scores.[20] The results for the other outcomes are equally small. We therefore conclude that it is unlikely that pre-existing family health trends – at least not those observable at birth – are responsible for our differential academic outcomes.

As a further test, in columns (2) and (4) of Table 3 we repeat the analysis from columns (1) and (3), but also control for birth weight, one and five-minute APGAR scores, and clinical estimate of gestational age at birth.[21] As can be seen in the table, the results are very similar to our baseline estimates. These findings provide additional evidence that factors that take place after birth and before the commencement of testing are responsible for the differences in birth order outcomes by third-born disability status. Sibling spillovers are a likely explanation.

We have also looked at other non-birth-outcome tests of unconfoundedness, and do not find statistically significant or large differences on other variables, and the patterns of (insignificant)

[19] In Table 4 for brevity we only present coefficients from the preferred specification where the third child is observed disabled by age five. Our results are unchanged if we focus on the other two specifications from Table 3: third observed disabled by age five or between ages five and ten; or third observed disabled by age ten. The results are also invariant to defining exposure based on spacing between siblings rather than simple birth order. In Appendix Table A1 we further present differences-in-differences comparisons akin to Table 2 for birth weight and five-minute APGAR score.

[20] If we apply the estimated relationships between birth weight and test scores reported by Figlio et al. (2014), the difference observed in birth weight in Table 4 would translate into test score differences of between 0.0009 and 0.0018 standard deviations, depending on the specification employed in Figlio et al.'s Table 2. Thus, the true relationship between birth weight and test scores would have to be 26 to 49 times larger in order to explain the test score differences we observe in Table 3 of this paper.

[21] As mentioned above, we do not observe birth outcomes for Danish children born before 1994, so we set these missing outcomes to zero and include a dummy for missing birth outcome (birth cohorts 1990 to 1993).The results of all columns in Table 3 are essentially unchanged if we restrict the Danish sample to children for whom we observe both birth and later life outcomes.

differences do not move in consistent ways – a finding suggesting that there exists no particular pattern of differential "trending" between families that would have a disabled third child versus those who would not have a disabled third child.[22] In order to test this more systematically, however, we estimate a model where we relate the disability status of the third child to the observable (and predetermined) characteristics of the earlier born children, an indicator for second-born, and the interaction of these characteristics with the second-born indicator. We then test whether the coefficients on the interaction terms are jointly equal to zero. These results are presented in Table 5, and suggest no differential trends in predetermined characteristics among families where the third child is observed with disability versus those where the third child is not observed with disability.

We conduct another test of unconfoundedness that we can carry out in Denmark but not in Florida due to data availability. We study the "effects" of a disabled third child on any interaction with hospitals of second-born versus first-born at age one and two – almost invariably *before the birth of the third child*. We have studied the "effects" on any contacts as well as contacts related to specific common diagnoses separately, and none of these is sizable or statistically significant. For instance, the "effect" on any interaction with the hospital system in the first year of life is 1.1 percentage points (standard error of 2.9), and mean of the dependent variable in this regression is 25 percent. In addition, we have studied mother's employment when the first two children are age 1 or 2, and again none of the "effects" are sizable or statistically significant. For example, when we measure maternal employment at age two (mean of 69.8 percent) the "effect" is 0.2 percentage points (standard error of 1.1).

We conduct another test of unconfoundedness that, this time, we can carry out in Florida but not in Denmark due to sample size constraints. Ideally, we would be able to study the "effects" of a disabled third child on the test score outcomes of second-born versus first-born children for families where the test scores are measured *before the third-born child is born*. Unfortunately, that would require families to have birth spacing of at least eight years between the second-born and third-born children in the case of Florida (and almost twice that amount of spacing in the case of Denmark), and only a handful of Florida families have this birth spacing pattern. However, we can

[22] When we examine time-varying individual characteristics, we find, for instance, that in Florida the difference between the second-born and the first-born siblings in the median income of the zip code at the time of birth is $353 (standard error of 293), or 2.5 percent of a standard deviation, lower when the third child is disabled versus when the third child is not disabled, and the difference in first grade school quality, based on measures by the state of Florida, is 0.36 percentile points (standard error of 0.29), or 2.2 percent of a standard deviation lower (See Autor et al. (2016) for a description of the percentiled school quality measure). On the other hand, Florida parents with a disabled third child are 1.3 percentage points (standard error of 1.0) more likely to get married between the first and second birth, relative to families without a disabled third child. In Denmark, families with a disabled third child are 0.3 percentage points (standard error of 0.5) less likely to be married or cohabiting at the time of the second birth than at the time of the first birth, relative to families without a disabled third child. The difference-in-differences in paternal income is small but statistically significant (-5223 Danish kroner, 3.3 percent of a standard deviation; standard error of 2007) but the difference-in-differences regarding maternal income is not statistically distinct from zero and is positively signed (809 kroner, 0.8 percent of a standard deviation; standard error of 1591). In sum, some patterns tend to have the first birth more advantaged in the case of a disabled third child, while other patterns tend to have the second birth more advantaged; none of the Florida patterns suggest statistically significant differences in these "pre-trends"; and the only statistically significant Danish difference is of such a small magnitude that it is highly unlikely to be driving the differences in sibling outcomes. Again, sibling spillovers are a more likely explanation for the differences in test scores.

look at an intermediate outcome – whether the second-born child is more likely to be identified as disabled by age five when the spacing between the second and third-born children in the family is at least five years. In this case, we have 372 Florida families where we can observe an outcome for the second-born child *before the third-born child is born.*[23] We find that second-born children who would eventually have disabled third-born siblings are slightly less likely to be identified as disabled before their third-born siblings are born. The point estimate is 0.8 percentage points less likely to be disabled than the first-born sibling with a standard error of 5.3, and the results are identical whether or not we control for birth outcomes. This biases us against finding evidence of sibling spillovers and suggests that our results are driven by actual sibling spillovers and not differential unobserved trajectories.[24]

Finally, because there is still the possibility that classification of the disability of the third child might be endogenously determined (if, say, the second-born sibling is performing poorly in school and that triggers increased vigilance and identification of the third-born as disabled), as an alternative to estimating the differential effects of third-sibling disability on second-born versus first-born siblings, we also consider the case in which the third-born sibling is observed with a significant issue at the time of birth as measured in the birth certificate – which we define as a congenital anomaly, an abnormal condition at birth, low birthweight (<2500g), prematurity (gestation less than or equal to 36 weeks), or poor health at birth (measured by five-minute APGAR scores below six on a ten-point scale). In Table 6, we present the results of this estimation.[25] We continue to find that if the third-born sibling has substantial issues at birth, the second-born sibling has worse schooling outcomes than does his/her first-born sibling. Second-born siblings score 4.5 percent of a standard deviation lower than do first-born siblings in Florida and 3.7 percent of a standard deviation lower in Denmark when the third-born sibling has substantial birth issues than when he/she does not. These results are again similar when we include birth certificate controls for the first- and second-born children. Thanks to this alternative variation we are less concerned that our results are somehow driven by endogenous identification of disability of the third-born sibling.

We have also attempted, in the case of Florida, to stratify disabilities on the basis of severity.

[23] We cannot apply this test in Denmark because there are only at most three families where the third-born child is disabled and the spacing between the second-born and third-born child is at least five years. This lack of variation is partially due to much lower mean disability, as defined in the two different data sources, in Denmark versus in Florida. Danish law prohibits mentioning sample information if the sample size is three or less.

[24] Consistent with a story of differential spillovers to the second-born versus first-born child, when the spacing between the second-born and third-born is much smaller, then the second-born is more likely to be identified as disabled. This last result could be due to a disabled third-born child differentially weakening the second-born child, or potentially due to endogenous identification of the third-born sibling's disability. To eliminate the potential concern about co-diagnosis of siblings, we estimate our models for families in which *neither* the first-born nor the second-born child is observed as disabled. In this case, we estimate that the (non-disabled) second-born child has scores that are 6.8 percent of a standard deviation lower (standard error of 3.1) in Florida and 4.3 percent of a standard deviation lower (standard error of 2.0) in Denmark than his/her (non-disabled) first-born sibling in the case in which the third-born sibling is observed disabled by age five. Again, the coefficients are nearly identical regardless of whether we include controls for birth outcomes.

[25] In this analysis we are able to utilize full sample for Denmark because all third-born children are born in 1994 or later when we can observe complete birth records. In column (4) of Table 6 we use the same imputation strategy as in columns (2) and (4) of Table 3. The results are unchanged if we restrict the sample to only individuals born after 1994. Abnormal conditions at birth are only observed in Florida.

We do not have airtight ways of measuring disability severity, but can proxy for severity based on the degree to which the state department of education compensates school districts for a student in a specific disability category. We find that the estimated effects of disabled third siblings are most negative when the third child has a moderately severe disability, relative to mild disabilities or very severe disabilities.[26] We believe that this makes sense because children with the most severe disabilities receive additional services through the state's Early Steps Program and other early childhood health and education support programs for disabled children, which might mitigate the effects of having a very disabled sibling, and those with mild disabilities might have less of an overall effect on the family. These different severity groups could therefore affect families differently due to how they influence parental time and financial resources available to older non-disabled siblings. In section 5 below we present some exploratory analysis that begins to investigate whether effects of disabled younger siblings on parental time and financial resources could be responsible for our results.

Finally, another concern might be that what we are picking up is not the effect of sibling disability but is in fact a result of differential family size due to endogenous fertility. To test this, we directly estimate the role of sibling disability on mother's subsequent fertility. When we examine the association between third child disability and the probability of a fourth child being born, we find no evidence of this. These results are presented in Appendix Table A6, and yield small, statistically insignificant and mixed-signed point estimates.

In summary, while we find sibling spillover effects using a variety of ways of measuring third-child disability, we do not observe evidence that these effects are the result of endogenously-determined disability identification. We further do not observe evidence that our results are due to co-diagnosis of siblings; we do not observe evidence that second-born children had any different outcomes observed before the third child was born (whether birth outcomes or early health problems), relative to first-born siblings, depending on third-sibling disability status; and we find no evidence that families whose third-born children are disabled were on any different trajectories in well-being prior to that birth than were families whose third-born children are not disabled. Moreover, our results are concentrated in the types of disabilities – physical disabilities and those observed early in life – where we would most expect to see differential spillovers on second-born versus first-born siblings. The results of these analyses provide some confidence that our empirical findings are not due to omitted variables that compromise the internal validity of our estimates.

5 Heterogeneity: Understanding mechanisms

We have found a general pattern, present in both Denmark and Florida, that a disabled third sibling differentially affects the second-born sibling relative to the oldest sibling. There are a number of plausible explanations for why this pattern would take place. For one, siblings who are more closely spaced in age are probably more likely to spend time together in mutual activities, so there are more

[26] The moderately severe disability group is also the largest, constituting 78 percent of all disabilities, while very severe disabilities are observed in only 3% of cases (42 families).

opportunities for direct spillovers between the third-born sibling and the second-born sibling than between the third-born sibling and the first-born sibling. But there is a set of indirect mechanisms through which this differential effect could take place as well. We think of these mechanisms broadly as involving time and money. On the time front, if a disabled younger sibling requires very time-intensive attention (e.g., additional medical visits, home-based therapies, and increased parental care) then there would be fewer opportunities for families to invest time in their older children. On the financial front, to the degree to which disabled siblings require costly therapies or equipment, or reduce parental labor supply (as is evidenced by the papers cited in footnote 2), there would be fewer financial resources available to invest in the older children. In both the financial and time resource case, the middle child should be more affected than the oldest child in the family, because the oldest child will necessarily be considerably older by the time the third sibling's disability becomes apparent and the family begins to make accommodations.

While we have no direct way of investigating the potential roles of time and financial resources in explaining our patterns of results, we believe that heterogeneity analysis may be helpful. For Denmark, we supplement these heterogeneity analyses with descriptive evidence, informative about health care utilization and mothers' use of the social services which may alleviate some of the constraints in terms of time and financial resources.[27]

We begin by exploring the differential effects by maternal education, and divide the samples into three groups, based on whether the mother is a four-year college graduate (24 percent in Florida, 28 percent in Denmark), a high school graduate or someone with some postsecondary education (55 percent in Florida, 49 percent in Denmark), or a high school dropout (or basic school completion only in Denmark; 21 percent in Florida, 23 percent in Denmark).[28] We suspect that the financial aspects of the sibling spillover may be more significant the less educated the mother, at least in the context of Florida with its less encompassing social safety net, given that many costs borne by families are similar across socio-economic status. On the time dimension, however, we suspect that the sibling spillover could have more substantial effects at the top of the SES distribution, given that highly-educated parents tend to spend considerably more time with their children than do less-educated parents (Guryan et al. 2008; Kalil et al. 2012), especially along quality-adjusted dimensions (Vinopal and Gershenson 2017).

Table 7 presents differences in the estimated effects of a disabled younger sibling by maternal education levels in Florida and Denmark. We observe that the patterns in the estimated relationships diverge across the two settings. In the case of Florida, there is an apparent U-shaped pattern, with the most negative estimated spillovers occurring at the bottom of the SES distribution (where financial resources are more likely to play a role) and at the top of the SES distribution (where time resources are more likely to play a role). In the Danish case, there is no apparent negative spillover

[27] We are unable to use the Danish Time Use Survey to directly address the time-constraint mechanism due to insufficient number of three-child families available in this data set.

[28] We code maternal education as the level of maternal education at the time of the first sibling's birth. We have also used it as an outcome and we do not observe any significant improvements in maternal education between first and second birth depending on the disability status of the third child. This provides another data point to our placebo analyses discussed in Section 4.1.

at the top of the SES distribution, but the negative spillover is present in the lower two groups of the SES distribution. It is important to note, however, that we cannot reject the null of equality of coefficients across education groups at conventional levels of statistical significance (the p-value of test of equality is 0.14 in Denmark and 0.45 in Florida).

Cultural and institutional factors might both play a role in determining which of these relationships may be more likely to dominate. The Danish and American social safety nets differ considerably, and the Danish welfare state offers heavily subsidized high-quality daycare including additional support for disabled children as well as income support programs for parents. According to the Law of Service, parents of severely disabled children can be employed by the municipality to care for their own children and can make claims for time lost if they are unable to work full time. Danish families have considerably more access to own sick leave, should the child disability affect parental mental health. In fact, we observe that mothers with a disabled third child are three percentage points less likely to work full time when the child is aged five and collect as much as 31% more in sickness benefits during the period from birth until the third child turns five (Table A3).[29] At the same time, they spend more time in doctor's offices and hospitals with their third-born children.

Denmark is also a much smaller country (around the size of Maryland) where people live closer to their extended families. The differences in proximity between young families and grandparents across the settings is obvious when we consider that only 34% of mothers of Florida-born children (and 22% of college-graduate mothers) were even born in Florida themselves, while in Denmark, 38% of families (26% for college-graduate mothers) reside in the same *municipality* as the child's maternal grandparents. And these municipalities are very small – on average 169 mi^2, 17% of the size of the average *county* in Florida.[30] Metropolitan Copenhagen alone consists of 34 municipalities within 1,170 mi^2, representing 36% of the nation's population, and the two largest Danish metropolitan areas combine to comprise 60% of the Danish population. New York's metropolitan area, in contrast, accounts for just 6% of the population of the United States, and one would have to add up the 18 largest American metropolitan areas (everything from Tampa-St. Petersburg, FL and larger) to account for the same share of the American population as metropolitan Copenhagen represents.

We suspect that the different pattern of results may reflect the fact that different options are available to cope with a disabled child in the two contexts. Appendix Table A2 shows, in fact, that the negative effects in Denmark are driven by parents living in different municipalities than the maternal grandmothers at the time of birth of the third child, which is consistent with the idea that parental and grandparental time with children may be relatively substitutable. Another potential explanation for the lack of a negative effect at the top of the SES distribution could be that these parents may be better at utilizing the public support available to cope with the challenges

[29] We do not find any differences in unemployment benefits utilization, and in fact mothers of disabled third-borns receive less leave benefits that are not directly related to their kids disability. We have also studied maternal labor supply and cohabitation status at age 10 as outcomes. The estimated effects are small and insignificant.

[30] The country is divided in 98 municipalities ranging from 3.5 mi^2 to 600 mi^2 (see www.noegletal.dk). In fact, 63% of the Danish population reside in municipalities smaller in area than the city of San Francisco, and 94% reside in municipalities smaller in area than New York City.

associated with raising a disabled child. We believe that these institutional differences explain why the magnitude of the overall effect is slightly smaller in Denmark compared to Florida, and why the pattern of heterogeneity across education of the mother varies across the two contexts.[31]

6 Conclusions

It is notoriously difficult to measure spillovers within a family, in part because finding and measuring an exogenous shock to family interaction is so difficult. We propose a strategy that allows us to causally identify a portion of the magnitude of a sibling spillover. Our insight is that the estimated effects of birth order should be different when the first- and the second-born siblings are differentially exposed to a family shock. We carry this out by studying the case of a disabled third sibling. Using detailed population-level data on thousands of three-plus-child families in Denmark and Florida, we observe consistent evidence indicating that the second-born child in a family is more adversely affected, relative to the first-born child, when the third child is disabled, a result that survives a battery of falsification tests. The magnitude of these differences is policy relevant; for example, in Florida the effect of the additional exposure to a disabled sibling that a second-born experiences relative to the first-born is about half of the observed relationship between an extra year of maternal education and children's test scores. We also provide some preliminary evidence that suggests that the sibling spillovers we measure work at least in part through relative constraints on parental time and financial resources. Our empirical approach could, of course, be applied to numerous other family shocks that affect one sibling differentially from another, and indeed, we find that the results are very similar if instead we study the differential effects of health problems of the third-born child observed at the time of birth.

While these results do not measure the effects of a specific policy, understanding sibling spillovers is important for helping to understand the full effects of family policy. For example, discussions concerning human capital interventions frequently consider what the benefits may be for the participating or targeted children themselves, but are less likely to involve thinking about the spillover effects on the family. Our results indicate that investments that ameliorate the effects of negative shocks to children likely have larger benefits than would be ascertained from considering the effects on the child alone. They may also have methodological implications for the interpretation of estimates based on sibling fixed effects identification strategies where one sibling is treated and the other forms a control group, and suggest that these may be downward biased.

[31] Stratification based on the immigrant status of the mother also supports this potential explanation as immigrants have access to fewer governmental resources. In both environments, we find estimates that are more negative (though not statistically significantly so) in the case of immigrant mothers than in the case of native-born mothers. In Florida, the differences-in-differences estimate is -0.066 (standard error of 0.066) for immigrant mothers and -0.042 (0.024) for native-born mothers. In Denmark, the differences-in-differences estimate is -0.119 (0.045) for immigrant mothers and -0.015 (0.020) for native-born mothers.

References

Adhvaryu, Achyuta and Anant Nyshadham, "Endowments at Birth and Parents' Investment in Children," *Economic Journal*, 2016, *126* (593), 781–820.

Alsan, Marcella, "The Gendered Spillover Effect of Young Children's Health on Human Capital: Evidence from Turkey," 2017.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman, "School Quality and the Gender Gap in Educational Achievement," *American Economic Review*, 2016, *106* (5), 289–295.

Bingley, Paul, Petter Lundborg, and Stepanie Vincent Lyk-Jensen, "Brothers in Arms: spillovers from a draft lottery," 2017.

Björklund, Anders and Markus Jäntti, "How Important is Family Background for Labor-Economic Outcomes?," *Labour Economics*, 2012, *19* (4), 465–474.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "The More The Merrier? The Effects of Family Size and Birth Order on Children's Education," *Quarterly Journal of Economics*, 2005, *120* (2), 669–700.

Booth, Alison and Hiau J. Kee, "Intergenerational Transmission of Fertility Patterns," *Oxford Bulletin of Economics and Statistics*, 2009, *71* (2), 183–208.

Breining, Sanni, "The presence of ADHD: Spillovers between siblings," *Economics Letters*, 2014, *124* (3), 469–473.

_ , Joseph Doyle, David Figlio, and Krzysztof Karbownik, "Birth Order and Delinquency: Evidence from Denmark and Florida," NBER Working Paper 23038, 2016.

_ , Meltem Daysal, Marianne Simonsen, and Mircea Trandafir, "Spillover Effects of Early Life Medical Interventions," IZA Discussion Paper 9086, 2016.

Brock, William and Steven Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies*, 2001, *68* (2), 235–260.

_ and _ , "Social Interactions and Technology Adoption," *Journal of the European Economic Association*, 2010, *8* (1), 232–251.

Burton, Peter, Lynn Lethbridge, and Shelley Phipps, "Children with Disabilities and Chronic Conditions and Longer-Term Parental Health," *Journal of Socio-Economics*, 2008, *37*, 1168–1186.

Conley, Dalton and Rebecca Glauber, "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility," *Journal of Human Resources*, 2006, *41* (4), 722–737.

Crowe, Terry, "Time Use of Mothers with Young Children: The Impact of a Child's Disability," *Developmental Medicine and Child Neurology*, 1993, *35* (7), 621–630.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

Currie, Janet and Mark Stabile, "Child Mental Health and Human Capital Accumulation: The Case of ADHD," *Journal of Health Economics*, 2006, *25* (6), 1094–1118.

Deshpande, Manasi, "Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls," *American Economic Review*, 2016, *106* (11), 3300–3330.

_ , "The Effect of Disability Payments on Household Earnings and Income: Evidence from the SSI Children's Program," *Review of Economics and Statistics*, 2016, *98* (4), 638–654.

Epple, Dennis and Richard Romano, "Peer Effects in Education: A Survey of the Theory and Evidence," in Jess Benhabib, Matthew Jackson, and Alberto Bisin, eds., *Handbook of Social Economics*, Elsevier, 2010.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth, "The Effects of Poor Neonatal Health on Children's Cognitive Development," NBER Working Paper 18846, 2013.

_ , _ , _ , and _ , "The Effects of Poor Neonatal Health on Children's Cognitive Development," *American Economic Review*, 2014, *104* (12), 3921–3955.

Fletcher, Jason and Barbara Wolfe, "Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited," *Journal of Health Economics*, 2008, *27* (3), 794–800.

_ , Nicole Hair, and Barbara Wolfe, "Am I my Brother's Keeper? Sibling Spillover Effects: The Case of Developmental Disabilities and Externalizing Behavior," NBER Working Paper 18279, 2012.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney, "Parental Education and Parental Time with Children," *Journal of Economic Perspectives*, 2008, *22* (3), 23–46.

Heckman, James, "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science*, 2006, *30* (5782), 1900–1902.

Joensen, Juanna Schroeter and Helena Skyt Nielsen, "Spillovers in Education Choice," *Journal of Public Economics*, 2018, *157*, 158–183.

Kalil, Ariel, Rebecca Ryan, and Michael Corey, "Diverging Destinies: Maternal Education and the Developmental Gradient in Time with Children," *Demography*, 2012, *49* (4), 1361–1383.

Kimmel, Jean, "Child Care Costs as a Barrier to Employment for Single and Married Mothers," *Review of Economics and Statistics*, 1998, *80* (2), 287–299.

Kvist, Anette, Helena Skyt Nielsen, and Marianne Simonsen, "The Importance of Children's ADHD for Parents' Relationship Stability and Labor Supply," *Social Science and Medicine*, 2013, *88*, 30–38.

Manski, Charles F., "Indentification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

_ , "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 2000, *14* (3), 115–136.

Mazumder, Bhashkar, "Sibling Similarities and Economic Inequality in the U.S.," *Journal of Population Economics*, 2008, *21* (3), 685–701.

Mitchell, Wendy, "Research Review: The Role of Grandparents in Intergenerational Support for Families with Disabled Children: A Review of the Literature," *Child and Family Social Work*, 2007, *12* (1), 94–101.

Moffitt, Robert A., "Models of Treatment Effects when Responses are Heterogeneous," *Proceedings of the National Academy of Science*, 1999, *96* (12), 6575–6576.

NCES, "Digest of Education Statistics, 2014," Technical Report 2016-006, National Center for Education Statistics 2016.

Nicoletti, Cheti and Birgitta Rabe, "Sibling Spillover Effects in School Achievement," IZA Discussion Paper 8615, 2014.

Noonan, Kelly, Nancy Reichman, and Hope Corman, "Mothers' Labor Supply in Fragile Families: The Role of Child Health," *Eastern Economics Journal*, 2005, *31* (4), 601–616.

Ozier, Owen, "Exploiting Externalities to Estimate the Long-term Effects of Early Childhood Deworming," *American Economic Journal: Applied Economics*, 2017, *forthcoming.*

Parman, John, "Childhood Health and Sibling Outcomes: Nurture Reinforcing Nature During the 1918 Influenza Pandemic," *Explorations in Economic History*, 2015, *58*, 22–43.

Powers, Elizabeth, "Children's Health and Maternal Work Activity: Estimates under Alternative Disability Definitions," *Journal of Human Resources*, 2003, *38* (3), 522–556.

Price, Joseph, "Parent-Child Quality Time: Does Birth Order Matter?," *Journal of Human Resources*, 2008, *43* (1), 240–265.

Qureshi, Javaeria, "Siblings, Teachers and Spillovers on Academic Achievement," *Journal of Human Resources*, 2017, *forthcoming.*

_ , "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital," *Economic Journal*, 2018, *forthcoming.*

Raaum, Oddbjørn, Kjell G. Salvanes, and Erik Sørensen, "The Neighbourhood is Not What it Used to be," *Economic Journal*, 2006, *116* (508), 200–222.

Reichman, Nancy, Hope Corman, and Kelly Noonan, "Effects of Child Health on Parents' Relationship Status," *Demography*, 2004, *41* (3), 569–584.

Salkever, David, "Children's Health Problems and Maternal Work Status," *Journal of Human Resources*, 1982, *17* (1), 94–109.

Seltzer, Marsha, Jan Greenberg, Frank Floyd, Yvette Pettee, and Jinkuk Hong, "Life Course Impacts of Parenting a Child with a Disability," *American Journal on Mental Retardation*, 2001, *106* (3), 265–286.

Solon, Gary, "Intergenerational Mobility in the Labor Market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3A, Elsevier, 1999, pp. 1761–1800.

Stabile, Mark and Sara Allin, "The Economic Costs of Childhood Disability," *The Future of Children*, 2012, *22* (1), 65–96.

Vinopal, Katie and Seth Gershenson, "Re-Conceptualizing Gaps by Socioeconomic Status in Parental Time with Children," *Social Indicators Research*, 2017, *forthcoming.*

Yi, Junjian, James Heckman, Junsen Zhang, and Gabriella Conti, "Early Health Shocks, Intra-Household Resource Allocation and Child Outcomes," *Economic Journal*, 2015, *125*, F347–F371.

Tables

Table 1: Descriptive statistics

	(1) 2+ children	(2) 3+ children	(3) 3+ children	(4) 3+ children
	All	All	3rd child observed disabled by age 5	3rd child not observed disabled by age 5
Panel A. Florida: Family characteristics				
African-American	0.12	0.17	0.14	0.18
	(0.32)	(0.38)	(0.35)	(0.38)
Hispanic	0.23	0.22	0.19	0.23
	(0.42)	(0.42)	(0.39)	(0.42)
Mother HS dropout	0.08	0.11	0.14	0.10
	(0.27)	(0.32)	(0.35)	(0.30)
Mother college graduate	0.33	0.29	0.26	0.30
	(0.47)	(0.46)	(0.44)	(0.46)
Mother's age at first birth	25.76	23.84	23.51	23.95
	(5.29)	(5.02)	(4.99)	(5.03)
Married at first birth	0.79	0.72	0.70	0.73
	(0.41)	(0.45)	(0.46)	(0.45)
Spacing 1st to 2nd	2.66	2.15	2.11	2.16
	(1.26)	(0.94)	(0.94)	(0.94)
Spacing 1st to 3rd	-	4.64	4.46	4.70
	-	(1.40)	(1.36)	(1.41)
Spacing 2nd to 3rd	-	2.49	2.35	2.54
	-	(1.15)	(1.08)	(1.17)
Number of families	80,879	9,987	2,483	7,504
Panel B. Denmark: Family characteristics				
Immigrant	0.11	0.16	0.17	0.16
	(0.31)	(0.36)	(0.38)	(0.36)
Mother HS dropout (or basic school only)	0.22	0.21	0.24	0.21
	(0.42)	(0.41)	(0.43)	(0.41)
Mother college graduate	0.24	0.26	0.24	0.26
	(0.43)	(0.44)	(0.43)	(0.44)
Mother's age at first birth	26.26	25.65	25.47	25.67
	(3.88)	(3.61)	(3.71)	(3.60)
Married/cohabiting at first birth	0.95	0.96	0.96	0.96
	(0.21)	(0.19)	(0.21)	(0.19)
Spacing 1st to 2nd	2.91	2.40	2.40	2.40
	(1.34)	(0.95)	(0.95)	(0.95)
Spacing 1st to 3rd	-	5.92	5.92	5.92
	-	(1.77)	(1.74)	(1.77)
Spacing 2nd to 3rd	-	3.48	3.48	3.48
	-	(1.56)	(1.57)	(1.55)
Number of families	134,277	28,581	1,882	26,699

Note: Only families with same father of first two children are included. The first column presents means in families with two or more children. The second column presents means in families with three or more children. The third column presents means in families with three or more children where the third child is observed disabled by age five. The fourth column presents means in families with three or more children where the third child is not observed disabled by age five.

Table 2: Difference-in-differences: Exposure to disabled younger sibling

	(1) More exposed (2nd)	(2) Less exposed (1st)	(3) Difference
Panel A. Florida: Test scores			
3rd observed disabled by age 5	0.156	0.268	-0.112***
	(0.885)	(0.900)	(0.015)
3rd not observed disabled by age 5	0.244	0.309	-0.064***
	(0.869)	(0.868)	(0.006)
Difference	-0.089***	-0.040***	-0.048**
	(0.011)	(0.011)	(0.023)
Number of observations (families)		104,130 (9,812)	
Panel B. Denmark: 9th grade gpa			
3rd observed disabled by age 5	-0.041	0.082	-0.123***
	(0.821)	(0.802)	(0.026)
3rd not observed disabled by age 5	0.029	0.118	-0.089***
	(0.806)	(0.813)	(0.007)
Difference	-0.070***	-0.036*	-0.034*
	(0.019)	(0.019)	(0.019)
Number of observations		57,162 (28,581)	

Note: This table presents the results of difference-in-differences models with no covariates included. Sample means and standard deviations are reported in the first two rows of columns (1) and (2) for both Florida and Denmark. T-test differences with standard errors in the first two rows of column (3) and first two columns of the third row for both Florida and Denmark. Differences-in-differences estimates are presented in the bottom right corner for each location. Outcome variables combine math and reading assessments. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 3: Main results: Effects of exposure to disabled younger sibling

	More exposed vs. less exposed		Exposure measured in years	
	(1)	(2)	(3)	(4)
Panel A. Florida: Test scores				
3rd observed disabled by age 5	-0.047**	-0.047**	-0.021**	-0.020**
	(0.023)	(0.023)	(0.010)	(0.010)
3rd observed disabled by age 5	-0.046**	-0.047**	-0.020**	-0.020**
	(0.023)	(0.023)	(0.010)	(0.010)
3rd observed disabled between age 5 and 10	0.008	0.005	0.007	0.005
	(0.026)	(0.026)	(0.011)	(0.011)
3rd observed disabled by age 10	-0.023	-0.024	-0.008	-0.009
	(0.018)	(0.018)	(0.008)	(0.008)
Controlling for birth outcomes		X		X
Mean of dependent variable		0.268		
Number of observations		104,130		
Number of families		9,812		
Panel B. Denmark: 9th grade gpa				
3rd observed disabled by age 5	-0.032*	-0.030	-0.010	-0.009
	(0.018)	(0.018)	(0.007)	(0.007)
3rd observed disabled by age 5	-0.033*	-0.030*	-0.010	-0.009
	(0.018)	(0.018)	(0.007)	(0.007)
3rd observed disabled between age 5 and 10	-0.017	-0.016	-0.010	-0.009
	(0.024)	(0.024)	(0.009)	(0.009)
3rd observed disabled by age 10	-0.027*	-0.025*	-0.010*	-0.009*
	(0.015)	(0.015)	(0.006)	(0.006)
Controlling for birth outcomes		X		X
Mean of dependent variable		0.070		
Number of observations		57,162		
Number of families		28,581		

Note: This table presents the results of difference-in-differences models with family fixed effects and controls for year and month of birth and the child's gender. Each panel has 12 regressions: four for the third child being observed disabled by age five; four for the third child being observed disabled by either age five or between ages five and ten; and four for the third child being observed disabled by age ten. Outcome variables combine math and reading assessments. Birth outcomes controls in columns (2) and (4) include infant birth weight, one and five minutes APGAR scores and clinical estimate of gestational age. Birth outcomes are not available in Denmark for birth cohorts 1990 to 1993. For these observation we impute zero in place of a birth outcome control and additionally include and indicator variable for missing value. Standard errors are adjusted for clustering at the family level. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 4: Testing unconfoundedness: Birth outcomes

	(1)	(2)	(3)	(4)
	Birth weight	Ln(birth weight)	Low birth weight	Apgar 5 score
Panel A. Florida: Birth outcomes				
3rd observed disabled by age 5	-16.591	-0.004	0.003	0.018
	(16.328)	(0.006)	(0.008)	(0.020)
Mean of dependent variable	3371	8.110	0.044	8.980
Number of observations		19,624		
Panel B. Denmark: Birth outcomes				
3rd observed disabled by age 5	-10.922	-0.004	0.002	-0.026
	(13.339)	(0.004)	(0.006)	(0.029)
Mean of dependent variable	3539	8.160	0.027	9.859
Number of observations		51,940		

Note: This table presents the results of difference-in-differences models with family fixed effects and controls for year and month of birth and the child's gender. Standard errors are adjusted for clustering at the family level. The Danish observations are smaller than in previous tables because we only observe birth outcomes for children born from 1994 forward, and thus for Denmark this table includes birth cohorts 1994 to 2001. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 5: Testing unconfoundedness: relationship between third child disability and predetermined characteristics

Second born interacted with time varying characteristic	(1)	(2)	(3)	(4)
	Third disabled by age 5		Third disabled by age 10	
	Florida	Denmark	Florida	Denmark
Married/cohabiting (A)	0.018**	0.009	0.013	0.008
	(0.009)	(0.011)	(0.011)	(0.013)
Age at birth (A)	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.000)
Education (A)	-0.001	0.000	-0.002	0.000
	(0.001)	(0.000)	(0.002)	(0.000)
Zip code/mother income (A)	-0.000	-0.009	0.000	0.000
	(0.000)	(0.015)	(0.000)	(0.019)
Health problems (A)	0.004	-0.002	-0.023	0.005
	(0.011)	(0.006)	(0.014)	(0.007)
Birth weight (B)	-0.000	-0.001	-0.000	-0.004
	(0.000)	(0.004)	(0.000)	(0.005)
Gestation weeks (B)	0.000	0.000	0.001	0.001
	(0.003)	(0.001)	(0.004)	(0.002)
Abnormal condition (B)	-0.005		-0.013	
	(0.022)		(0.027)	
Congenital anomalies (B)	0.035	0.035	0.072	0.019
	(0.073)	(0.026)	(0.085)	(0.030)
F-test A (p-value)	0.32	0.62	0.35	0.37
F-test B (p-value)	0.64	0.60	0.61	0.78
F-test A & B (p-value)	0.59	0.71	0.42	0.62
Mean of Y	0.137	0.067	0.243	0.104
Observations	18,915	51,489	18,915	51,489

Note: This table regresses an indicator for third child disability observed by age 5 (columns 1 and 2) or by age 10 (columns 3 and 4) on a set of interactions between second born and predetermined (to third birth) characteristics of mother (A) or first two children (B), levels of these characteristics and second born indicator. Table displays the interactions. Family fixed effects are not included because the outcome does not vary within mother, however, standard errors are adjusted for clustering at the family level as there are two observations per family in the analysis sample. Maternal characteristics include: indicator for being married (FL) or married/cohabiting (DK); age at the time of birth, years of education, zip code of residence income (FL) or mother's income first year after child's birth (DK); health problems as recorder on birth certificate (FL) or maternal interaction with hospital second year after child's birth (DK). Maternal interactions with hospital are not available for births cohorts 1990 to 1993, and we include indicator for missing observations as well as its interaction with second born (not displayed). This interaction is included when performing F-test. Children characteristics include: birth weight, gestational age, abnormal conditions at birth (only FL) as well as congenital anomalies. F-tests at the bottom of the table test hypotheses that maternal characteristics (A), child characteristics (B) or both sets are jointly equal to zero. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 6: Using health issues at the time of birth as an alternative to measured disability

	(1)	(2)	(3)	(4)
	Florida		Denmark	
3rd observed with birth problems	-0.045*	-0.041*	-0.037**	-0.032*
	(0.025)	(0.025)	(0.017)	(0.017)
Controlling for birth outcomes		X		X
Mean of dependent variable	0.268		0.070	
Number of observations	104,130		57,162	
Number of families	9,812		28,581	

Note: Difference-in-differences models with family fixed effects and controls for year and month of birth and the child's gender. Outcome variables combine math and reading assessments. Treatment is defined as following indicators recorded on birth certificate: congenital anomaly or abnormal condition at birth (only FL) or birth weight below 2500 grams or gestational age below 37 weeks or five minutes Apgar score below 6. Standard errors clustered at the family level. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table 7: Heterogeneity by maternal education

	(1) College graduate	(2) Complete HS or with some college	(3) HS dropout (or basic school only)
Panel A. Florida: Test scores			
3rd observed disabled by age 5	-0.078*	-0.021	-0.080
	(0.045)	(0.030)	(0.054)
Mean of dependent variable	0.814	0.217	-0.257
Number of observations	25,680	57,256	21,194
Number of families	2,387	5,373	2,052
Panel B. Denmark: 9th grade gpa			
3rd observed disabled by age 5	0.030	-0.058**	-0.031
	(0.036)	(0.027)	(0.040)
Mean of dependent variable	0.484	0.121	-0.391
Number of observations	14,706	25,616	12,206
Number of families	7,353	12,808	6,103

Note: This table presents the results of difference-in-differences models with family fixed effects and controls for year and month of birth and the child's gender. Maternal education is measured at the time of first birth. Outcome variables combine math and reading assessments. Standard errors are adjusted for clustering at the family level. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Appendix Tables

Table A1: Difference-in-differences: Unconfoundedness

	(1) More exposed (2nd)	(2) Less exposed (1st)	(3) Difference
Panel A. Florida			
Birth weight			
3rd observed disabled by age 5	3411	3338	73***
	(531)	(564)	(21)
3rd not observed disabled by age 5	3417	3324	93***
	(501)	(518)	(8)
Difference	-6	14	-20
	(15)	(15)	(17)
Five minutes Apgar score			
3rd observed disabled by age 5	9.00	8.96	0.04**
	(0.42)	(0.54)	(0.02)
3rd not observed disabled by age 5	8.99	8.97	0.02***
	(0.47)	(0.55)	(0.01)
Difference	0.01	-0.01	0.02
	(0.01)	(0.02)	(0.02)
Number of observations		19,624	
Panel B. Denmark:			
Birth weight			
3rd observed disabled by age 5	3600	3440	160***
	(544)	(532)	(18)
3rd not observed disabled by age 5	3625	3456	169***
	(512)	(510)	(5)
Difference	-25**	-16	-9
	(13)	(13)	(13)
Five minutes Apgar score			
3rd observed disabled by age 5	9.87	9.84	0.03
	(0.82)	(0.82)	(0.03)
3rd not observed disabled by age 5	9.89	9.83	0.06***
	(0.75)	(0.82)	(0.01)
Difference	-0.02	0.01	-0.03
	(0.02)	(0.02)	(0.02)
Number of observations		51,940	

Note: Difference-in-differences models with no covariates included. Sample means and standard deviations are reported in the first two rows of columns (1) and (2) for both Florida and Denmark. T-test differences with standard errors in the first two rows of column (3) and first two columns of the third row for both Florida and Denmark for each outcome. Differences-in-differences estimates are presented in the bottom right corner for each location and outcome. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A2: Heterogeneity in the effects by access to grandparents in Denmark

	(1) Living in the same municipality as grandmother	(2) Living in a different municipality than grandmother
3rd observed disabled by age 5	0.025	-0.046*
	(0.033)	(0.026)
Mean of dependent variable	0.003	0.227
Number of observations	17,034	28,288
Number of families	8,517	14,144

Note: Difference-in-differences models with family fixed effects and controls for year and month of birth and the child's gender. Outcome variables combine math and reading assessments. We define the grandmother as the mother of the child's mother. We only consider families where the grandmother is not dead at the time of birth of the third-born, and thus assign grandparental living status based on third birth. Standard errors clustered at the family level. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A3: Utilization of welfare and health care systems

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Welfare system (mothers)				Health care utilization (child)		
	Employed	Leave benefits	Unemployment benefits	Sickness benefits	Inpatient admissions	Outpatient admissions	ER admissions
3rd observed disabled by age 5	0.724	29351	42544	20390	3.691	2.791	1.069
	(0.447)	(44565)	(76641)	(57585)	(4.566)	(2.731)	(1.494)
Observations	1,862		1,697			1,882	
3rd not observed disabled by age 5	0.753	33504	42657	15589	1.556	0.472	0.776
	(0.431)	(47057)	(78621)	(48686)	(1.964)	(0.995)	(1.166)
Observations	25,996		23,705			26,699	
Difference (t-test)	-0.029***	-4153***	-113	4801***	2.135***	2.319***	0.293***
	(0.010)	(1178)	(1972)	(1240)	(0.053)	(0.028)	(0.028)
Observations	27,858		25,402			28,581	

Note: This table presents utilization of welfare system by mothers and health care system by a third-born child in Denmark. Unit of analysis is a household. For all outcomes it shows means and standard deviations, measured for third birth, for families where the third child is disabled by age 5 and for families where the third child is not disabled by age 5. The last last row presents difference in means and standard errors based on t-test. The outcome variables are: probability that mothers is employed full-time when third child is 5 years old (column 1), amount of leave benefits received (column 2), amount of unemployment benefits received (column 3), amount of sickness benefits received (column 4), number of inpatient admissions (column 5), number of outpatient admissions (column 6) and number of emergency room admissions (column 6). Columns (2) to (4) present cumulative amounts from one to five years after the birth of the third child in DKK. Leave benefits in column (2) include child care leave (one additional year on top of maternity leave), education leave (time off to improve one's education) and sabbatical (unrestricted one year off work). Columns (5) to (7) present counts of the events by age five. Differences marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A4: Difference-in-differences: Main results by disability type

	(1)	(2)	(3)	(4)
	Florida		Denmark	
	Physical	Cognitive or behavioral	Physical	Cognitive or behavioral
3rd observed disabled by age 5	-0.073***	0.012	-0.043**	0.011
	(0.027)	(0.040)	(0.020)	(0.066)
3rd observed disabled by age 5	-0.075***	0.014	-0.044**	0.011
	(0.027)	(0.040)	(0.020)	(0.066)
3rd observed disabled between ages 5 and 10	-0.031	0.022	-0.026	-0.002
	(0.035)	(0.037)	(0.033)	(0.038)
3rd observed disabled by age 10	-0.060***	0.018	-0.039**	0.002
	(0.022)	(0.028)	(0.017)	(0.033)
Mean of dependent variable	0.314	0.255	0.074	0.078
Number of observations	92,716	89,184	55,604	49,466
Number of families	8,719	8,409	27,802	24,733

Note: This table replicates analysis presented in columns (1) and (2) of Table 3 separately for treatment defined as physical disabilities (columns 1 and 3) and cognitive or behavioral disabilities (columns 2 and 4). Standard errors are adjusted for clustering at the family level. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A5: Difference-in-differences: Main results for mathematics and reading

	Mathematics		Reading	
	(1)	(2)	(3)	(4)
Panel A. Florida: Test scores				
3rd observed disabled by age 5	-0.031	-0.031	-0.062**	-0.061**
	(0.024)	(0.024)	(0.024)	(0.024)
3rd observed disabled by age 5	-0.030	-0.030	-0.061**	-0.061**
	(0.025)	(0.025)	(0.024)	(0.024)
3rd observed disabled between age 5 and 10	0.010	0.009	0.008	0.004
	(0.028)	(0.028)	(0.028)	(0.028)
3rd observed disabled by age 10	-0.012	-0.013	-0.031	-0.033*
	(0.020)	(0.020)	(0.019)	(0.019)
Mean of dependent variable	0.283		0.253	
Number of observations	102,714		103,814	
Number of families	9,805		9,811	
Panel B. Denmark: 9th grade gpa				
3rd observed disabled by age 5	-0.032	-0.030	-0.010	-0.008
	(0.023)	(0.023)	(0.029)	(0.029)
3rd observed disabled by age 5	-0.032	-0.030	-0.009	-0.007
	(0.023)	(0.023)	(0.029)	(0.029)
3rd observed disabled between age 5 and 10	0.002	0.005	0.015	0.016
	(0.030)	(0.030)	(0.038)	(0.038)
3rd observed disabled by age 10	-0.019	-0.017	-0.000	0.001
	(0.019)	(0.018)	(0.023)	(0.023)
Mean of dependent variable	0.124		0.052	
Number of observations	55,966		55,658	
Number of families	28,546		28,505	
Controlling for birth outcomes		X		X

Note: This table replicates analysis presented in columns (1) and (2) of Table 3 separately for mathematics (columns 1 and 2) and reading (columns 3 and 4). Standard errors are adjusted for clustering at the family level. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.

Table A6: First-difference models: subsequent fertility

	(1)	(2)	(3)
	Denmark		Florida
	4th child born within 5 years	4th child born within 10 years	4th child born within the sample
3rd observed disabled by age 5	-0.006	0.005	0.010
	(0.006)	(0.007)	(0.008)
3rd observed disabled by age 5	-0.006	0.006	0.010
	(0.006)	(0.007)	(0.008)
3rd observed disabled between age 5 and 10	-0.003	0.006	0.005
	(0.008)	(0.010)	(0.009)
3rd observed disabled by age 10	-0.005	0.006	0.008
	(0.005)	(0.006)	(0.007)
Mean of dependent variable	0.181	0.249	0.086
Number of families	28,581		9,812

Note: This table relates fertility decisions following third child and third child disability status, where disability is defined as observed by age five, between ages five and ten or by age ten. There are 9 regressions in this table, three for each definition of treatment. Columns (1) and (2) present results for Denmark where we can observe completed fertility while column (3) presents results for Florida where we can only observe subsequent births up to birth cohort 2002. Regressions include only third child's birth year fixed effects as additional controls, and there is one observation per family in the analysis sample. Heteroskedasticity robust standard errors. Estimates marked ***, **, and * are statistically significant at the 1, 5, and 10 percent levels, respectively.