THE ECONOMIC JOURNAL

The Economic Journal, 1–28 DOI: 10.1093/ej/ueaa074

SIBLING SPILLOVERS*

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

It is notoriously difficult to identify peer effects within the family. Using administrative data on children from both Florida and Denmark, the paper examines the effects of having a disabled younger sibling. To address the identification challenge, the paper compares the differential effects for first- and second-born children in three-plus-child families, taking advantage of the fact that birth order influences the amount of time that a child spends in early childhood with their younger siblings, disabled or not. The paper finds evidence that, relative to the first born, the second child in a family is differentially affected when the third child is disabled.

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 behaviours. 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 how they operate, the presence of spillovers across siblings implies that policies or shocks that affect one child in a family may have important external effects on other children, suggesting a potential for either positive or negative multiplier effects.

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, 2011).

* Corresponding author: Sandra E. Black, Department of Economics, Columbia University, 1022 International Affairs Building, 420 West 118th Street, New York, NY 10027, USA. Email: sblack@columbia.edu

This paper was received on 25 May 2018 and accepted on 8 June 2020. The Editor was Frederic Vermeulen.

The authors were granted an exemption to publish their data because access to the data is restricted. However, the authors provided a simulated or synthetic dataset, which enabled the Journal to run their codes. The synthetic/simulated data and codes are available on the Journal website. They were checked for their ability to generate all tables and figures in the paper, but the synthetic/simulated data are not designed to reproduce the same results.

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, Michael Lovenheim, Richard Romano, and seminar and conference participants at AEA, AEFP, APPAM, CALDER, CESifo, NBER 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.

[1] As we discuss later, notable exceptions include Fletcher *et al.* (2012), Breining (2014), Parman (2015), Yi *et al.* (2015), Adhvaryu and Nyshadham (2016), Breining *et al.* (2015), Alsan (2017), Heissel (2017), Joensen and Nielsen (2018), Ozier (2018), Qureshi (2018a), Qureshi (2018b), Bingley *et al.* (2020), Karbownik and Özek (2019) and Nicoletti and Rabe (2019).

For this reason, 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 to 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. Even though our estimation is based on a specific situation, we believe it allows us to learn about how families function more broadly. Children with disabilities require different kinds of investments of time, attention and financial resources from parents; any resulting reallocation of these inputs may affect the development of other children in the family.[2] 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 school year 2012-13, for instance, in the United States alone over 6.4 million children aged 3 to 21 (12.9% 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 biases associated with unobserved time-invariant family characteristics, and isolate the effects of siblings using a difference-in-differences, within-family comparison. While this approach allows us to estimate the effects of relative exposure to a disabled sibling, we are unable to estimate the total effect of exposure to a disabled sibling. Because families with a disabled child may be different in unobserved

[2] An extensive literature documents the consequences of child disability for parents. The increased complexity of parenting children with disabilities affects parental labour 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), 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' labour 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.

[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; Black *et al.*, 2018; Breining *et al.*, 2020).

ways from other families, we are unable to distinguish whether differences in outcomes are the result of exposure to a disabled child or of differences in unobservables across these families. Another limitation of our research design is that it requires us to restrict our analysis to three-child families. While three-child families are not representative of all families, we believe it is valuable to address the difficult internal validity challenges inherent to peer effects estimation even if it comes at the expense of external validity to some degree. It is worth noting, however, that because our design measures the effect of relative exposure, and not the total effect of having a disabled sibling, our estimates may understate the total effect of having a disabled younger sibling.[4]

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 and early postnatal 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 or early postnatal 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. We find they are not. 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 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 are considered disabled by age 5 in families with disabled and non-disabled third children. We find that first- and second-born siblings are equally likely to be considered disabled by age 5 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 disabilities diagnosed later 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 data sets where we merge full birth cohorts with schooling and medical 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

[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. Two recent papers—Fletcher *et al.* (2012), using Panel Study of Income Dynamics-Child Development Supplement (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.

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 school 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 standardised 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.)[5]

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-standardised grade from the 9th 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, healthcare systems and social welfare systems. To the degree to which we find comparable results in these two quite different environments, it further increases our confidence that the estimates are externally valid and more generally reflect spillovers and dynamics within the family, as opposed to 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 significant; 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 driven by physical disabilities—which are likely to be more visible early and require more parental time and attention—rather than cognitive or behavioural disabilities.

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,

[5] We pooled a child's FCAT observations across all grades (3–8) to reduce the potential consequences of measurement error. We do require, however, that each pair of children is observed at least in grade 3 and in every subsequent grade utilised in the empirical sample. That said, our results are stable if we focus on a certain grade as opposed to pooling across multiple grades, but in some grades these estimates are no longer statistically significant at conventional levels due to increased standard errors in the smaller samples. This is due to the fact that youngest cohorts in our sample are not old enough to reach middle school by the time our data end. We observe test scores for all cohorts up to grades 4 or 5, depending on month of birth with respect to school entry cutoff of 1 September.

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 through strain on parental time and financial resources.

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 have 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 within family dynamics. Some of those have generated evidence of sibling effects in developing countries or within the United States a century ago.[6]

More directly related to our work, other studies in developed countries have examined the effects of educational investments in older children on their younger siblings. These studies have tended to focus on changes that occurred during pre-adolescence or adolescence, and to measure the effect of those investments on educational outcomes of younger siblings. The timing of the shocks to the older siblings in these studies suggests it is plausible they measure the direct influence older siblings' behaviour has on younger siblings' behaviour 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. In this context Landersø *et al.* (2020) and Karbownik and Özek (2019) would be closest to our current work in that they explicitly consider parental reallocation of resources as one of the channels but do so in an educational rather than health shock setting.[7]

[6] Adhvaryu and Nyshadham (2016) evaluate an iodine supplementation programme in Tanzania and find that there are effects on the treated children's siblings as well as the target child. Alsan (2017) examines an immunisation campaign in Turkey and finds positive effects on the human capital of children not age-eligible for the programme but whose siblings were eligible. Ozier (2018) evaluates a deworming intervention in Kenya and finds positive test score effects on siblings of children who were treated through the programme. Qureshi (2018a) uses gender segregation of schools in Pakistan to measure the effect of older siblings' schooling on their younger siblings, finding that increases in older sisters' 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 parental reallocation of resources can be a mechanism by which one sibling's health affects 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.

[7] Karbownik and Özek (2019) use regression discontinuity design to study sibling spillovers in academic achievement in Florida. They find that younger siblings of old-for-grade students, as compared to those of young-for-grade students, have higher test scores but these effects are present only among resource-constrained households. In more affluent households, they find some evidence for parental reinforcement if younger sibling in a family is unusually high performer due to their school entry age. Landersø *et al.* (2020) use that same econometric design and policy variation but examine only spillovers running from a younger to an older siblings in Denmark. They find that being older at school start increases maternal employment and marital stability, however, effects on older sibling achievement are much more heterogeneous and present only when the older child is at the crucial stage of their educational career (9th grade exit exams). Joensen and Nielsen (2018) study siblings in Denmark to estimate effects of an older sibling taking an advanced maths–science course during high school on their younger siblings. They find that it increases the likelihood their younger sibling will take an advanced maths–science course when they reach high school by 2–3 percentage points. Using data from North Carolina, Qureshi (2018b) 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 (2019) 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.* (2020) 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%; while Heissel (2017) studies sibling spillovers from teenage pregnancies in Florida and shows that both brothers and sisters of teen mothers are negatively affected.

Our work is one of the few papers, in a developed country context, to examine the effect of a health shock to a child during early developmental years on the educational outcomes of other siblings; 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 on 9th grade exit exams, and Breining *et al.* (2015), 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 maths 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 other 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.

1. 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.[8]

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—due to data availability—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. 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, labour market status, earnings, leave taking and immigrant status.

We measure disability differently in the two settings, and we provide detailed description of the conditions that form our treatment in Online Appendix A. In Florida, disabilities are recorded on the school records in mutually exclusive categories. 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. The most common disabilities first observed by age 5 in Florida are speech impairment (48%), developmental delay (21%) and language impairment (17%). The most common disabilities first observed between ages 5 and 10 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 5 in Denmark are congenital malformations and deformations

[8] 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.

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 5 and 10 are behavioural and emotional disorders (29%), disorders of psychological development (14%) and congenital malformations of eye, ear, face and neck (8%).[9]

Our identification strategy requires us to observe at least three births in a family. Furthermore, because it is helpful for conditions to be 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 on a larger number of families in Denmark, despite the fact that Florida is more populous.[10] 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 5 implying much higher disability diagnoses rate in Florida. This discrepancy is likely due to the fact that we observe a broader range of disability types in Florida, as can be seen in Online Appendix A.

Table 1 presents descriptive statistics for our analysis population.[11] 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 families with two children and the same father for the first two children, while column 2 shows descriptives for 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 of families with two children is demographically diverse. In the Florida sample, 11% of the mothers are African-American, and 23% are Hispanic.[12] Of sampled Florida mothers, 7% are high school dropouts, and 34% 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.[13] The Danish sample is less racially and ethnically diverse than the Florida population, though 9% of the mothers in families with two children in Denmark are immigrants. Denmark also has more high school dropout (22%) and fewer college graduate (23%) mothers in the analysis sample. Just as in Florida, the Danish sibling sample is comparatively privileged in terms of socio-economic background.[14] The age at first birth, however, is quite similar on average: 26.0 in Florida versus 26.4 in Denmark.

The sample of families with at least three children is somewhat different from the sample of families with only two children (conditional on the first two having the same father in both

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

[10] 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. In our empirical sample, since the birth order is sourced from birth certificates, we can observe it without measurement error net of the usual possibility of coding errors in administrative data.

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

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

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

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

Table 1. *Descriptive Statistics.*

	(1) 2 children All	(2) 3+ children All	(3) *t*-test (1) vs. (2)	(4) 3+ children 3rd child observed disabled by age 5	(5) 3+ children 3rd child not observed disabled by age 5	(6) *t*-test (4) vs. (5)
Panel A: Florida: family characteristics						
African-American	0.11	0.17	−18.3	0.14	0.18	4.1
	(0.31)	(0.38)		(0.35)	(0.38)	
Hispanic	0.23	0.22	1.4	0.19	0.23	4.7
	(0.42)	(0.42)		(0.39)	(0.42)	
Mother HS dropout	0.07	0.11	−14	0.14	0.10	−4.8
	(0.26)	(0.32)		(0.35)	(0.30)	
Mother college graduate	0.34	0.29	8.8	0.26	0.30	3.9
	(0.47)	(0.46)		(0.44)	(0.46)	
Mother's age at first birth	26.03	23.84	39.1	23.51	23.95	3.8
	(5.27)	(5.02)		(4.99)	(5.03)	
Married at first birth	0.80	0.72	17.6	0.70	0.73	2.6
	(0.40)	(0.45)		(0.46)	(0.45)	
Spacing 1st to 2nd	2.73	2.15	43.9	2.11	2.16	1.9
	(1.29)	(0.94)		(0.94)	(0.94)	
Spacing 1st to 3rd	–	4.64	–	4.46	4.70	7.4
		(1.40)		(1.36)	(1.41)	
Spacing 2nd to 3rd	–	2.49	–	2.35	2.54	7.4
		(1.15)		(1.08)	(1.17)	
Number of families	70,892	9,987	80,879	2,483	7,504	9,987
Panel B: Denmark: family characteristics						
Immigrant	0.09	0.16	−31.6	0.17	0.16	−1.4
	(0.29)	(0.36)		(0.38)	(0.36)	
Mother HS dropout (or basic school only)	0.22	0.21	3.8	0.24	0.21	−2.7
	(0.42)	(0.41)		(0.43)	(0.41)	
Mother college graduate	0.23	0.26	−8.5	0.24	0.26	2.0
	(0.42)	(0.44)		(0.43)	(0.44)	
Mother's age at first birth	26.43	25.65	29.9	25.47	25.67	2.1
	(3.94)	(3.61)		(3.71)	(3.60)	
Married/cohabiting at first birth	0.95	0.96	−6.2	0.96	0.96	1.3
	(0.21)	(0.19)		(0.21)	(0.19)	
Spacing 1st to 2nd	3.05	2.40	74.5	2.40	2.40	−0.1
	(1.39)	(0.95)		(0.95)	(0.95)	
Spacing 1st to 3rd	–	5.92	–	5.92	5.92	−0.1
		(1.77)		(1.74)	(1.77)	
Spacing 2nd to 3rd	–	3.48	–	3.48	3.48	0.0
		(1.56)		(1.57)	(1.55)	
Number of families	105,696	28,581	134,277	1,882	26,699	28,581

Notes: Only families with same father of first two children are included. Columns (1), (2), (4) and (5) present means and standard deviations (in parentheses). Column (1) is restricted to families with two children; column (2) is restricted to families with three or more children; column (4) is restricted to subset of families from column (2) where third child is observed with disability by age 5; while column (5) is restricted to subset of families from column (2) where third child is not observed with disability by age 5. Column (3) presents *t*-test statistics for differences between columns (1) and (2) while column (6) presents *t*-test statistics for differences between columns (4) and (5).

samples). In Florida, families with at least three children are more likely to have African-American mothers (17% versus 11% for two-child families), but have similar shares of Hispanic mothers. In Denmark, families with three or more children are more likely to have immigrant mothers (16% versus 9%). In Florida, mothers with less education are more likely to have families with three or more children (11% versus 7% high school dropout mothers, 29% versus 34% college graduate mothers), though the pattern is reversed in Denmark (26% versus 23% college graduate mothers).

Partly because all Florida births in the data occurred within a nine-year window, the average spacing between first and third births in Florida (4.6 years) is smaller 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 more likely to have first- and second-born children who are born closer together. On average, among those with exactly two children, the second-born child arrives when the first-born is 2.7 years old in Florida and 3.1 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 4 and 5 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 to be African-American or Hispanic. 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% versus 73%). These differences are generally smaller in Denmark as compared to Florida which could reflect the different definitions of disability. Irrespective of the exact cause, these differences—albeit mostly modest in magnitude—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 a research design that relies on comparisons within families and that controls for differences across families.

2. Difference-in-Differences Empirical Strategy

Our research approach is to carry out a simple difference-in-differences design: we compare second- versus first-born children, in families with disabled and non-disabled third children. The basic idea underlying the research design is that, because we observe outcomes at a fixed common age, second-born children have spent a larger share of their lives to that point exposed to a disabled sibling than first-borns have. 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 differences in unobservables. Our research design accounts for those differences across families by making comparisons within families, across first- and second-born

siblings, who have different amounts of exposure to the third-born siblings by the time they reach any given age.

Our main regression equation is the following:

$$Y_{if} = \alpha_f + \beta_0 + \beta_1 M_{if} + \beta_2 D_f M_{if} + \beta_3 \mathbf{X}_{if} + \varepsilon_{if} \quad i = 1, 2, \tag{1}$$

where Y denotes a student's standardised test scores or grade point average (normalised to have mean zero and standard deviation one), 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, $\mathbf{X}$, includes indicator variables for year and month of birth as well as child gender. The index i denotes individuals and f denotes family, and α_f is 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 any time invariant differences between families with and without a disabled third child. This is an estimate specific to families with three or more children, and as demonstrated in Table 1 these families are somewhat different in their observable characteristics from families with two children—the minimal family size for spillovers to exist. Although we do not have a causal design at hand that would allow estimation of these effects in smaller families, we attempt to test for generalisability by re-estimating our models for families with three or more children weighted with characteristics of households with two children only. These results are qualitatively similar to our main findings.

The key identifying assumption in a difference-in-differences approach is that there are no differential trends between the treatment and control group in the absence of the 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 the outcomes of children in families with disabled third children were already on a downward trajectory 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.[15] We investigate this common trends assumption by studying, among others, 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 (1) for the entire sample as well as for several subgroups defined by, e.g., mother's education, type of disability or its severity. 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 of variation in birth spacing instead of birth order. In Subsection 3.1 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.

[15] 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, as documented later, we do not observe that third-child disability is correlated with either subsequent fertility or family size.

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 (families)		57,162 (28,581)	

Notes: 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. Difference-in-differences estimates are presented in the bottom right corner of each panel. Sample sizes reflect children–year observations in Florida (grades 3–8) and children observations in Denmark (grade 9 only) while in parentheses we provide number of families for these children. Since in Florida our children–year observations are not a balanced panel we observe some children in all 3–8 grades while other children in only subset of those grades. Thus, average grade for Florida children in this data set is 5.3 and median is 5. In Denmark we only observe achievement in a single 9th grade. Outcome variables combine maths 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% levels, respectively.

3. 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 based on the categorisation in school records at kindergarten entry, around age 5. As can be seen in the first and second column, first-born children score higher on standardised tests in 3rd through 8th grade than do second-born children; this could be for many reasons, such as differential exposure to undivided parental attention, 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. These birth order differences are comparable to what has been documented in the literature using all first- and second-born children (Breining *et al.*, 2020). 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 3rd–8th 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 at conventional levels ($p = 0.073$). The smaller estimated effects for Denmark may reflect the fact that the educational, healthcare and social welfare systems are focused on alleviating the impact of disadvantages such as disabilities on the individual and the family. It could also be a consequence of the fact that we pool grades 3–8 in Florida and focus on grade 9 only in Denmark, but our grade 8 estimate for Florida is −0.041 (SE of 0.032)—exactly in between the preferred estimates in the two locations. Meanwhile, our estimate for grade 3 is −0.052 (SE of 0.026)—slightly larger than that from the last grade.

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 5 on a second-born sibling's outcomes relative to the first-born sibling's outcomes. These results are very 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 by the time of outcome measurement which is at a fixed age. 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 Column 4 in Table 3, and point towards the same conclusion.

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 5 versus the case in which the third child is identified as disabled at some point between ages 5 and 10. The purpose of this specification is to gauge the degree to which relatively late-observed disabilities appear to differentially affect second-born versus first-born siblings. The results suggest that the effect of a disabled third sibling appears to be larger when the disability is observed earlier. 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 5 and 10, 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

Table 3. *Main Results: Effects of Exposure to Disabled Younger Sibling.*

	More exposed vs. less exposed			Exposure measured in years		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Florida: test scores						
3rd observed disabled by age 5	−0.047**	−0.047**	−0.048**	−0.021**	−0.020**	−0.021**
	(0.023)	(0.023)	(0.023)	(0.010)	(0.010)	(0.010)
3rd observed disabled by age 5	−0.046**	−0.047**	−0.047**	−0.020**	−0.020**	−0.020**
	(0.023)	(0.023)	(0.023)	(0.010)	(0.010)	(0.010)
3rd observed disabled between age 5 and 10	0.008	0.005	0.004	0.007	0.005	0.005
	(0.026)	(0.026)	(0.026)	(0.011)	(0.011)	(0.011)
3rd observed disabled by age 10	−0.023	−0.024	−0.025	−0.008	−0.009	−0.009
	(0.018)	(0.018)	(0.018)	(0.008)	(0.008)	(0.008)
Birth outcomes		X	X		X	X
Marital status and income			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	−0.031*	−0.010	−0.009	−0.009
	(0.018)	(0.018)	(0.018)	(0.007)	(0.007)	(0.007)
3rd observed disabled by age 5	−0.033*	−0.030*	−0.031*	−0.010	−0.009	−0.010
	(0.018)	(0.018)	(0.018)	(0.007)	(0.007)	(0.007)
3rd observed disabled between age 5 and 10	−0.017	−0.016	−0.016	−0.010	−0.009	−0.009
	(0.024)	(0.024)	(0.024)	(0.009)	(0.009)	(0.009)
3rd observed disabled by age 10	−0.027*	−0.025*	−0.026*	−0.010*	−0.009*	−0.010*
	(0.015)	(0.015)	(0.015)	(0.006)	(0.006)	(0.006)
Birth outcomes		X	X		X	X
Marital status and income			X			X
Mean of dependent variable			0.070			
Number of observations			57,162			
Number of families			28,581			

Notes: 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 18 regressions: six for the third child being observed disabled by age 5 (first row); six for the third child being observed disabled by either age 5 or between ages 5 and 10 (next two rows); and six for the third child being observed disabled by age 10 (last row). Each regression is using the same set of observations and only first two births are included in the estimation sample. Treatment is defined based on disability status of the third child, and third-borns are not included in the estimation sample. In the first row, treatment group is families where third-born child is observed with disability by age five while control group is families where third-born child is not observed with disability by age five. In the next two rows, the treatment groups are families where third-born child is observed with disability by age 5 or between ages 5 and 10 while control group is families where third-born child is not observed with disability by age 10. Finally in the last row of regression estimates, the treatment group is families where third-born child is observed with disability by age 10 while control group is families where third-born child is not observed with disability by age 10. Dashed lines denote different regression sets. Each column is a different set of regressions: three for Florida and three for Denmark. Columns (1)–(3) interact disability with birth order while columns (4)–(6) with birth spacing. Birth outcomes controls in columns (2) and (5) include infant birth weight, five minutes Apgar score indicators and weeks of gestational age indicators. Birth outcomes are not available in Denmark for 9.1% of observations. For these observation we impute zero in place of a birth outcome control and additionally include and indicator variable for missing value. Additional socio-economic controls in columns (3) and (6) include marital status and combined parental income (zip code level neighbourhood income in Florida) at birth. Outcome variables combine maths 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% levels, respectively.

cases in which the third-born child was observed disabled by age 10; this is simply a combination of those first observed disabled by age 5 and those first observed disabled between ages 5 and 10. 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 are comparatively young (Heckman, 2006; Cunha and Heckman, 2007).[16] Panel B shows the parallel results for the Denmark sample, which are smaller in magnitude and less precisely estimated, but are similar in pattern to the results in Florida.[17]

Our preferred estimates of sibling spillovers, where the third child is identified as disabled by age 5 (column 1), are 4.7% and 3.2% of a standard deviation for Florida and Denmark, respectively. To assess the magnitude of these estimates, it is helpful to compare them to other estimates in the family, education, and health literature, and in particular to those obtained using data from these same sources. For instance, the Florida effect size is about the same as the effect of a 10% increase in birth weight (Figlio *et al.*, 2014), 80% of the birth order gap in reading scores (Breining *et al.*, 2020), a quarter of the size of the effect of school entry cutoff on cognitive development (Dhuey *et al.*, 2019), and four-fifths of the spillover effect of an older sibling's increased education (Karbownik and Özek, 2019). Likewise, the Danish effect size is about two-thirds of the effect of a 10% increase in birth weight (Kreiner and Sivertsen, 2018), or a quarter of the birth order gap in test scores (Breining *et al.*, 2020). Furthermore, Landersø *et al.* (2020) estimates 16% of a standard deviation gain among older siblings if their younger sibling is old for grade at the time of their elementary school exit exam. This reduced-form effect is over five times larger than our preferred estimate here.

Another intuitive benchmark for our causal estimates is the correlation between disability and own test scores. Pooling the data across first three births we estimate associations of −0.27 SD (SE of 0.02) and −0.10 SD (SE of 0.01) between test scores/grades and own disability for Florida and Denmark, respectively. These coefficients are larger for cognitive and behavioural as compared to physical disabilities, a distinction that we come back to below, but even in the latter case the estimated spillovers are only about 30% and 50% of this association in Florida and Denmark, respectively.[18]

[16] In the Danish data we can observe hospitalisation records, and those disabled by age 5 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 5 and 10.

[17] Descriptive statistics presented in Section 1 suggest that families with at least three children, that we need to execute our empirical design, are somewhat different from families with two children—minimal family size for spillovers to exist. Therefore, to gauge the extent to which our results are sensitive to these differences, we re-estimated models from Table 3 weighted with characteristics of households with two children only. These results are presented in Online Appendix Table B1, and point to similar—albeit statistically insignificant due to increased standard errors—negative effects in our preferred specification measuring disability by age 5. This exercise again suggests our results may be more broadly generalisable.

[18] It is also useful to compare these spillover effects with achievement gaps between demographic groups observed in the data. For example, the mean test score difference between students with high school dropout (or basic school only) and college graduate mothers is 110% and 55% of a standard deviation in Florida and Denmark, respectively. Thus, our preferred estimates constitute 4.3% and 5.8% of these gaps. Likewise, differences in test scores between children of married and unmarried parents are 58% and 27% of a standard deviation in the two locations; implying spillover effects of about 8.1% and 11.9% of these differences, respectively. More generally, our more negative spillover effect of almost 5% of a standard deviation is equivalent to a decrease of $1,000 in Earned Income Tax Credit (EITC) income or the difference in the average EITC payout today versus that in the early 1990s (Dahl and Lochner, 2012); or one-quarter of the effects of assignment of a small class for grades K–3 on test scores or one-half of the effect of the same assignment on ACT/SAT performance (Krueger and Schanzenbach, 2001); or at least 10% of the effect of childcare subsidies on children's GPA (Black *et al.*, 2014). They are also comparable to the effects of parental job loss, which is consistent with dominating indirect rather than direct effect; however, in such a case parental financial resources decrease but parental time with children may actually increase (although the quality of this time is uncertain). For example, using Norwegian data, Rege *et al.* (2011) estimate detrimental effects of about 5% to 6% of a standard deviation which is similar to our main effects. Thus, we view these magnitudes as not only plausible but also comparable to findings in other studies of the determinants of cognitive development.

In addition, we have examined the effects of particular groups of disabilities (see Online Appendix Table B2). In particular, we compare the spillover effects of physical versus cognitive and behavioural disabilities. We only find deleterious effects in the case of physical disabilities, which constitute about 55% and 76% of all disability diagnoses of third-borns in Florida and Denmark, respectively. This may suggest that the main mechanism through which the spillovers operate is a tightening of household budget constraints—e.g., financial or parental time— rather than through direct interactions between siblings.[19] For example, using Danish data, we observe that relative to healthy third-born children, excess rates of outpatient and inpatient admissions are twice as large for children with physical disabilities as for children with cognitive or behavioural disabilities (see Online Appendix Table B3). This suggests that the time constraint is strained more by the former rather than the latter set of conditions.

We also examine whether or not there are differences in reading versus mathematics. The results, shown in Online Appendix Table B4, indicate larger effects on reading than 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.

3.1. *Testing for Unconfoundedness*

We conduct a number of robustness and specification checks to test our key identifying assumptions. First, it may be the case that families with disabled third children were trending in a manner that was different from those without disabled third children even prior to the birth of the third child. If children in families with disabled third children were becoming progressively comparatively less healthy or academically successful for reasons that were 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 first consider four different measures of two birth outcomes observed consistently in both Denmark and Florida: birth weight in grams, log of birth weight, low birth weight (defined as birth weight under 2,500 g), and five-minute Apgar scores (observed on a ten-point scale). We first focus on the pre-trends in health rather than socio-economic outcomes because our shock—disability—may be more directly related to the former rather than the latter set of circumstances. In fact, as explained later in this section, we can use poor birth outcomes of the third child as an alternative treatment variable to estimate sibling spillovers and in such context birth outcomes of the older siblings should provide relevant evidence on pre-trends in children's health. Table 4 presents the analogous analysis to Table 3, but with these dependent variables that were established prior to the birth of the third sibling, showing the coefficients on the interaction between the third sibling's disability status by age 5 and second-born status in the

[19] It could also be that older siblings need to care for their younger disabled brother or sister, however, we expect such an effect to be present across various disabilities and not limited to only physical impairments. It is also plausible that physical disability like a speech impairment affect third-born child's communication skills and educational outcomes which could have a direct effect on other siblings in the family. In our view such effect would have to operate through detrimental interactions between siblings rather than direct transmission of human capital and cognitive ability as younger children will generally, by the virtue of birth ordering, possess less skill and knowledge on average, and therefore are less likely to teach or mentor older siblings. This, in fact, is the exact logic behind findings in Qureshi (2018b) and Nicoletti and Rabe (2019) that highlight direct channels as dominant.

Table 4. *Testing Unconfoundedness: Birth Outcomes.*

	(1) Birth weight	(2) ln(birth weight)	(3) Low birth weight	(4) 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	3,371	8.109	0.044	8.980
Number of observations		19,624		19,614
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	3,539	8.160	0.027	9.859
Number of observations		51,940		

Notes: 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. The number of observations for Denmark is smaller than in previous tables because birth outcomes are missing for 9.1% of observations. Standard errors are adjusted for clustering at the family level. Estimates marked ***, ** and * are statistically significant at the 1%, 5% and 10% levels, respectively.

family in a model with family fixed effects and indicators for gender, year and month of birth.[20] If the identifying assumptions for the sibling spillover are satisfied, these models should show no effect on outcomes that were predetermined at the time of the third child's birth. 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. In Online Appendix Table B6 we further present difference-in-differences comparisons akin to Table 2 for birth weight and five-minute Apgar score. Consistent with earlier research, we find that second-born children are heavier than first-borns; with statistically significant birth weight differences at 73 to 93 g or 160 to 169 g in Florida and Denmark, respectively. At the same time, consistent with the key identifying assumptions for the sibling spillover estimates, we find no statistically significant difference in these patterns among families with a disabled third sibling relative to those without; the relative differences due to the 'exposure' to a disabled third-born child are much smaller and statistically insignificant at 20 g and 9 g in Florida and Denmark, respectively.

Importantly, we find no apparent relationship between third-born disability and the differential birth outcomes of second- versus first-born siblings (Table 4).[21] 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 5) third-born

[20] In Table 4 for brevity we only present coefficients from the preferred specification where the third child is observed disabled by age 5. Our results are unchanged if we focus on the other two specifications from Table 3: third observed disabled by age 5 or between ages 5 and 10; or third observed disabled by age 10. The results are also invariant to defining exposure based on spacing between siblings rather than simple birth order. We report these additional results in Online Appendix Table B5, where out of 48 estimates only three are marginally statistically significant at 10% level. In both Denmark and Florida we now have only a single observation for each child and two children per family in the estimation sample.

[21] It is also conceivable that our small average birth effects mask substantial heterogeneity across groups that simply offset each other when pooled together. We test this possibility in Online Appendix Table B7, which replicates results from Table 4 but includes interactions between our treatment variable and various sociodemographic characteristics of the household. Irrespective of the sample and outcome we cannot reject the hypothesis that our effects are homogeneous, and from among 48 interactions only two are statistically significant at conventional levels. Furthermore, the effect sizes for our preferred neonatal health measure, natural logarithm of birth weight, never exceed 3%.

Table 5. *Testing Unconfoundedness: Time-Varying Characteristics.*

	(1)	(2) Florida	(3)	(4)	(5) Denmark	(6)
	Median zip code income	1st grade school quality	Parents married at birth	Paternal income	Maternal income	Parents married or cohabiting at birth
3rd observed disabled by age 5	−353	−0.362	0.013	−6,020**	−1,728	0.003
	(293)	(0.285)	(0.010)	(2,567)	(1,972)	(0.006)
Mean of dependent variable	46,919	53.857	0.776	216,840	109,135	0.964
SD of dependent variable	13,879	15.876	0.417	155,920	107,075	0.187
Number of observations	18,915	18,471	19,624	53,076	53,130	53,130
Number of families	9,679	9,611	9,812	26,538	26,565	26,565

Notes: 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. Dependent variables are household characteristics that vary across births. Columns 1–3 present outcomes for Florida while columns 4–5 present outcomes for Denmark. Column 1 is median zip code income measured at the time of birth; column 2 is school quality in 1st grade (Autor *et al.,* 2016); column 3 is an indicator for parents being married at the time of birth; columns 4 and 5 are paternal and maternal incomes measured at the calendar year in the year before the child's birth; and column 6 is an indicator for parents married or cohabiting and the time of birth. Standard errors are adjusted for clustering at the family level. Estimates marked ***, ** and * are statistically significant at the 1%, 5% and 10% levels, respectively.

sibling, 17 g in Florida or 11 g in Denmark, these differences are tiny relative to the estimated effects of a disabled third-born sibling on differential test scores.[22] The results for the other outcomes are equally small, and thus we 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.[23]

We have also looked at other non-birth-outcome tests of unconfoundedness, and do not find statistically significant or large differences on other variables either, and the signs of (insignificant) differences do not move in consistent ways—a finding suggesting that there is 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. We present these results in Table 5. When we examine time-varying individual characteristics, we find 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, or 2.5% of a standard deviation, lower when the third child is disabled versus when the third child is not disabled, and the difference in 1st grade school quality, based on

[22] 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.

[23] Online Appendix Tables B8 and B9 document parallel analyses for other birth outcomes that are either location specific (e.g., we can only observe one-minute Apgar score in Florida while leave benefits in Denmark) or available for limited samples (e.g., congenital anomalies in Denmark). Consistent with the necessary identifying assumptions for the spillover estimates, irrespective of the sample and location we do not find any evidence for differential health of the first two births by disability status of the third-born child. These estimates are statistically insignificant and small in magnitude. In each case we have also tested for the presence of heterogeneous treatment effects, and we reject them at conventional levels in 11 out of 13 specifications. In the two cases for Denmark where we cannot reject the null hypothesis, for gestational age and leave benefits, the average effect sizes are 0.05% and 8%, respectively; controlling for these time-varying controls does not alter our main results.

measures by the state of Florida, is 0.36 percentile points, or 2.3% of a standard deviation lower (see Autor *et al.*, 2016 for a description of the percentiled school quality measure).[24] On the other hand, Florida parents with a disabled third child are 1.3 percentage points more likely to get married between the first and second birth, relative to families without a disabled third child. Likewise, in Denmark, families with a disabled third child are 0.3 percentage points more 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 (−6,020 Danish kroner, or 3.9% of a standard deviation) but the difference-in-differences regarding maternal income is not statistically distinct from zero (−1,728 kroner, or 1.6% of a standard deviation). In sum, along some dimensions, the first born is relatively more advantaged in families with a disabled third born, and along other dimensions, the second born is relatively more advantaged in these families. 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.

We also 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 6, 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.

As a further test, in columns (2), (3), (5) and (6) of Table 3 we repeat the analysis from columns (1) and (4), but also control for either only birth weight, five-minute Apgar scores and clinical estimate of gestational age at birth (columns 2 and 5) or those birth covariates together with marital status and income that also vary across births (columns 3 and 6).[25] *Ex ante* inclusion of these variables as controls could matter for two reasons. First, to the extent that they are correlated both with treatment and some third unobserved factors that affect test scores, they may account for this correlation and change the estimated coefficient of interest. Second, this alternative test addresses the possibility that our balancing tests lack precision to detect pre-trends. As can be seen in the table, the results are almost identical to our baseline estimates irrespective of which exact set of time-varying controls we include. 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.

As another test of unconfoundedness that we can carry out in Denmark but not in Florida, we study the 'effects' of a disabled third child on differential use of hospitals by first- and second-borns during the ages of 1 and 2—almost invariably *before the birth of the third child*. We have estimated the 'effects' on any hospital use as well as hospital use related to specific common diagnoses separately, and none of these is statistically significant. We present these results in

[24] School quality analysis should be treated with caution here because for majority of the families in the sample it is measured after the birth of third child, and thus can be thought of as endogenous to third-child's disability labelling. Nonetheless, it is encouraging that this time-varying and predictive of test scores measure appears balance between first- and second-born children in families with and without disabled third child.

[25] As mentioned above, we do not observe birth outcomes for 9.1% of Danish children, so we set these missing outcomes to zero and include a dummy for missing birth outcome. 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.

Table 6. *Testing Unconfoundedness: Relationship between Third-Child Disability and Predetermined Characteristics.*

Second born interacted with time-varying characteristic	(1) Third disabled by age 5 Florida	(2) Third disabled by age 5 Denmark	(4) Third disabled by age 10 Florida	(5) Third disabled by age 10 Denmark
Married/cohabiting (A)	0.017*	0.017	0.012	0.014
	(0.009)	(0.011)	(0.011)	(0.014)
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	0.004
	(0.000)	(0.016)	(0.000)	(0.019)
Father income (A)		−0.008		−0.013
		(0.008)		(0.010)
Health problems (A)	0.005	−0.002	−0.022	0.005
	(0.011)	(0.006)	(0.014)	(0.007)
Birth weight (B)	−0.000	−0.001	−0.000	−0.003
	(0.000)	(0.004)	(0.000)	(0.004)
Gestation weeks (B)	−0.000	0.000	0.000	0.001
	(0.003)	(0.001)	(0.004)	(0.002)
Five minutes Apgar (B)	0.011	−0.003	0.024**	−0.003
	(0.009)	(0.003)	(0.012)	(0.004)
Abnormal conditions (B)	−0.002		−0.006	
	(0.022)		(0.027)	
Congenital anomalies (B)	0.035	0.033	0.071	0.016
	(0.073)	(0.026)	(0.085)	(0.030)
F-test A (*p*-value)	0.55	0.38	0.18	0.27
F-test B (*p*-value)	0.35	0.58	0.39	0.83
F-test A & B (*p*-value)	0.54	0.50	0.21	0.54
Mean of *Y*	0.138	0.067	0.244	0.104
Observations	18,905	51,128	18,905	51,128

Notes: This table regresses an indicator for third-child disability observed by age 5 (columns 1–2) or by age 10 (columns 3–4) on a set of interactions between second-born and predetermined (to third birth) time-varying 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/family 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 the year before the calendar year of the child's birth (DK) and father's income measured the year before the calendar year of child's birth (DK); health problems as recorded 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 1990–3, 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. Child's characteristics include: birth weight, gestational age, five-minutes Apgar score, 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% levels, respectively.

Online Appendix Table B10. In addition, we have included mother's employment when the first two children are age 1 or 2 as a dependent variable, and again none of the 'effects' are sizable or statistically significant. For example, when we measure maternal employment at age 2 (mean of 69.8%) the 'effect' is 0.2 percentage points. We present these results in Online Appendix Table B11.

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, because the earliest we observe test scores is at age 8, 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 look at an intermediate outcome—whether the second-born child is more likely to be identified as disabled by age 5 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*.[26] 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 in this regression 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 estimate is statistically insignificant and the sign of the point estimate biases us against finding evidence of sibling spillovers and suggests that our results are driven by actual sibling spillovers and not differential unobserved trajectories within families.[27]

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 issues 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 (<2,500 g), prematurity (gestation fewer than or equal to 36 weeks) or poor health at birth (measured by five-minute Apgar scores below six on a ten-point scale). Note that Tables 4, B8 and B9 documented no differential pre-trends in these variables across the first- and second-born children and as evidenced below our results are not sensitive to inclusion of birth outcomes for the first two births in this analysis. In Table 7, we present the results of this estimation. 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. The gap between second-born and first-born siblings is 4.5% and 3.7% of a standard deviation in Florida and Denmark, respectively, when the third child has serious birth issues compared to when he/she does not. These results are again similar when we include birth certificate controls for the first- and second-born children. These findings point against the main spillover estimates being driven by endogenous identification of disability of the third-born sibling. Likewise, it is worth highlighting that this is a meaningful alternative

[26] 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 fewer.

[27] 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, the second-born is more likely to be identified as disabled. This 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 re-estimated 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% of a standard deviation lower in Florida and 4.3% of a standard deviation lower in Denmark than his/her (non-disabled) first-born sibling in the case in which the third-born sibling is observed disabled by age 5. Again, the coefficients are nearly identical regardless of whether we include controls for birth outcomes. We present these results in Online Appendix Table B12.

Table 7. *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	

Notes: Difference-in-differences models with family fixed effects and controls for year and month of birth and the child's gender. Outcome variables combine maths 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 2,500 g or gestational age below 37 weeks or five minutes Apgar score below 6. Standard errors are adjusted for clustering at the family level. Estimates marked ***, ** and * are statistically significant at the 1%, 5% and 10% levels, respectively.

specification, because although poor birth outcomes predict childhood disability in the statistical sense this correlation is only 0.03.

We have also attempted, in the case of Florida, to stratify disabilities on the basis of severity. We do not have an airtight way of measuring disability severity, but can proxy for it based on the degree to which the state department of education compensates school districts for a student in a specific disability category. Our logic here is that disabilities for which schools receive higher compensation are likely more time and resource demanding, and thus more severe from the perspective of not only educators but also parents. These results are presented in Online Appendix Table B13, and we find that the estimated effects of disabled third-born siblings are most negative when the third child has a moderately severe disability, relative to mild disabilities or very severe disabilities.[28]

This pattern in the results may be because children with the most severe disabilities receive additional services through the Florida's Early Steps Program and other early childhood health and education support programmes 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 4 we present some exploratory analysis that begins to investigate whether effects of disabled younger siblings on parental time and financial resources are responsible for our results.

Finally, another concern might be that what we are picking up is not the effect of sibling disability but rather 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

[28] Mild disabilities group includes only specific learning disability; moderate disabilities group includes educable mentally handicapped, trainable mentally handicapped, speech impaired, language impaired, deaf or hard of hearing, emotionally handicapped and intellectual disability; while severe disabilities include orthopedically impaired, visually impaired, severely emotionally disturbed, hospital/homebound, profoundly mentally handicapped, dual sensory impaired, autistic, traumatic brain injury and developmentally delayed. The moderately severe disability group is also the largest, constituting 78% of all disabilities, while severe disabilities are observed in only 3% of cases (42 families). Very small number of families experiencing severe disabilities could be another reason for why we do not find meaningful estimates in Online Appendix Table B13 for this group.

find no evidence of a relationship. These results are presented in Online Appendix Table B14, 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, various tests point against these effects being the result of endogenously determined disability identification. We also do not observe evidence that our results are due to co-diagnosis of siblings; we do not observe evidence of differential gaps between first- and second-born children before the third child was born (whether in birth outcomes or early health problems), 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, as noted in Section 3, 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.

4. 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 child relative to the oldest child. There are a number of plausible explanations for why this pattern would exist. 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 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. That the results are only present for physical disabilities which are diagnosed very early in childhood suggests that the direct channel may not be the primary driver of the sibling spillover we estimate. Were it due to direct effects stemming from engagement with the disabled sibling, we would have expected the effects to be present across a range of disabilities and ages.

At the same time, 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 extent disabled siblings require costly therapies or equipment, or reduce parental labour 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 about healthcare utilisation and mothers' use of the social services which may alleviate some of the constraints in terms of time and financial resources.

We begin by exploring differential effects by maternal education and divide the samples into three groups, based on whether the mother is a four-year college graduate (24% in Florida, 28%

Table 8. *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

Notes: 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 maths 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% levels, respectively.

in Denmark), a high school graduate or has some postsecondary education (55% in Florida, 49% in Denmark), or a high school dropout (or basic school completion only in Denmark; 21% in Florida, 23% in Denmark). We code maternal education based on the level of maternal education at the time of the first sibling's birth. 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 8 presents differences in the estimated effects of a disabled younger sibling by maternal education levels in Florida and Denmark. We observe that the estimated relationships are different 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). These two coefficients are about four times the size of the estimate for middle educational group. Likewise, the proposed U-shaped pattern holds only for physical disabilities, as documented in Online Appendix Table B15, where we expect smaller direct spillovers. In the Danish case, there is no apparent negative spillover 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 this evidence should be treated as suggestive, since we cannot reject the null of equality of coefficients across education groups at conventional levels of statistical significance (the *p*-values of test of equality are 0.14 in Denmark and 0.45 in Florida).

Cultural and institutional factors might play a role in determining which of these relationships—time or money—is dominant. The Danish and American social safety nets differ considerably, and the Danish welfare state offers heavily subsidised high-quality day care including additional

support for disabled children as well as income support programmes 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 sick leave, should the child disability affect parental mental health. In fact, we observe that mothers with a disabled third child are 3 percentage points less likely to work full time when the child is aged 5, and collect as much as 31% more in sickness benefits during the period from birth until the third child turns 5 (Online Appendix Table B16).[29] At the same time, we do not find any differences in unemployment benefits utilisation, and in fact mothers of disabled third-borns receive less leave benefits that are not directly related to their child's disability. Finally, as documented in columns 5–7 of Online Appendix Table B16, these families spend more time in doctor's offices and hospitals with their third-born children. Overall, we view this descriptive evidence as suggestive of the strain that a disabled third-born child can put on parental resources, at least in terms of time. Likewise, it is plausible that our estimates are larger in Florida than Denmark because parents in Florida, without such a broad access to the safety net, face both time and financial constraints. Although these correlations are only suggestive of certain pathways, we view them as potentially promising and policy relevant avenues for future research.

Denmark is also a comparatively small 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 33% 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 geographically small—17% of the size of the average *county* in Florida.[30] Metropolitan Copenhagen alone consists of 34 municipalities, 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.

Based on these geographic features, we suspect that the relative size of the estimated spillover effects may reflect the fact that different options are available to help families with a disabled child in the two contexts. Online Appendix Table B17 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 or even more so by families where grandmaternal care is not available at all, which is the case if the grandmother is dead or if the mother is an immigrant. This pattern of results is consistent with the idea that parental and grandparental time with children may be relatively substitutable, and that the spillover effect of a disabled sibling in part operates through a strain on parental time available to the non-disabled siblings. Another potential explanation for the lack of a negative effect at the top of the SES distribution could

[29] We have also investigated maternal and paternal labour supply using our preferred difference-in-differences strategy. In this case we define dependent variable as parental employment by age 5 of either first- or second-born child. The results indicate no economically meaningful or statistically significant differential effects of third-child's disability on parental employment in this setting. Note that such an analysis does not preclude the level differences described above.

[30] The country is divided in 98 municipalities ranging from 3.5 mi^2 to 600 mi^2 (see http://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.

be that these parents may be better at utilising the public support available to cope with the challenges associated with raising a disabled child. This could be either due to differential access and take up of services or because high-SES parents are more efficient in utilising available services. Since we cannot offer any direct evidence on either of these channels we treat our heterogeneity analyses as suggestive, and a valuable avenue for future research.[31]

We believe that these institutional differences could help in explaining 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. Stratification based on the immigrant status of the mother, presented in Online Appendix Table B18, 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 difference-in-differences estimate is −0.066 for immigrant mothers and −0.042 for native-born mothers. In Denmark, the contrast is more striking: difference-in-differences estimate is −0.119 for immigrant mothers and −0.015 for native-born mothers.

5. 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. We measure disability differently in the two settings, supporting the notion that our results reflect household dynamics more broadly.

The magnitude of the sibling spillover effects we measure is significant; 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 to suggest 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 non-disability health problems of the third-born child observed at the time of birth.

While the sibling spillovers we measure do not measure the effects of a specific policy, understanding sibling spillovers is important for helping to understand the full effects of family

[31] To get some additional purchase on these issues we have also analysed employment and welfare utilisation at third birth (Online Appendix Table B16) as well as the role of maternal grandmother presence (Online Appendix Table B17) separately by maternal education levels. We find that employment differences and welfare utilisation differences compared across the disability status of third child are smallest in families where mother has college degree. Likewise, these are the families where we do not find negative effects of maternal grandmother absence. Both these results are consistent with null effects on test scores in families with college educated mothers (Table 8) and could be due to the fact that these affluent families could afford to buy help and additional resources on the free market.

policies and shocks. 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 fully consider 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. Our findings 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.

Columbia University
Aarhus University
Northwestern University
Northwestern University
Emory University
Aarhus University
University of Florida
Aarhus University

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Adhvaryu, A. and Nyshadham, A. (2016). 'Endowments at birth and parents' investment in children', Economic Journal, vol. 126(593), pp. 781–820.
Alsan, M. (2017). 'The gendered spillover effect of young children's health on human capital: evidence from Turkey', NBER Working Paper 23702.
Autor, D., Figlio, D., Karbownik, K., Roth, J. and Wasserman, M. (2016). 'School quality and the gender gap in educational achievement', *American Economic Review*, vol. 106(5), pp. 289–95.
Bingley, P., Lundborg, P. and Vincent Lyk-Jensen, S. (2020). 'Brothers in arms: spillovers from a draft lottery', *Journal of Human Resources*, forthcoming.
Björklund, A. and Jäntti, M. (2012). 'How important is family background for labor–economic outcomes?', *Labour Economics*, vol. 19(4), pp. 465–74.
Black, S., Devereux, P., Løken, K. and Salvanes, K. (2014). 'Care of cash? The effect of child care subsidies on student performance', *Review of Economics and Statistics*, vol. 96(5), pp. 824–37.
Black, S., Grönqvist, E. and Öckert, B. (2018). 'Born to lead? The effect of birth order on noncognitive abilities', *Review of Economics and Statistics*, vol. 100(2), pp. 274–86.
Black, S.E., Devereux, P.J. and Salvanes, K.G. (2005). 'The more the merrier? The effects of family size and birth order on children's education', *Quarterly Journal of Economics*, vol. 120(2), pp. 669–700.
Booth, A. and Kee, H.J. (2009). 'Intergenerational transmission of fertility patterns', *Oxford Bulletin of Economics and Statistics*, vol. 71(2), pp. 183–208.
Breining, S. (2014). 'The presence of ADHD: spillovers between siblings', *Economics Letters*, vol. 124(3), pp. 469–73.
Breining, S., Daysal, M., Simonsen, M. and Trandafir, M. (2015). 'Spillover effects of early life medical interventions', IZA Discussion Paper 9086.
Breining, S., Doyle, J., Figlio, D., Karbownik, K. and Roth, J. (2020). 'Birth order and delinquency: evidence from Denmark and Florida', *Journal of Labor Economics*, vol. 38(1), pp. 95–142.
Brock, W. and Durlauf, S. (2001). 'Discrete choice with social interactions', *Review of Economic Studies*, vol. 68(2), pp. 235–60.
Brock, W. and Durlauf, S. (2010). 'Social interactions and technology adoption', *Journal of the European Economic Association*, vol. 8(1), pp. 232–51.
Burton, P., Lethbridge, L. and Phipps, S. (2008). 'Children with disabilities and chronic conditions and longer-term parental health', *Journal of Socio-Economics*, vol. 37, pp. 1168–86.

Conley, D. and Glauber, R. (2006). '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*, vol. 41(4), pp. 722–37.
Crowe, T. (1993). 'Time use of mothers with young children: the impact of a child's disability', *Developmental Medicine and Child Neurology*, vol. 35(7), pp. 621–30.
Cunha, F. and Heckman, J. (2007). 'The technology of skill formation', *American Economic Review*, vol. 97(2), pp. 31–47.
Currie, J. and Stabile, M. (2006). 'Child mental health and human capital accumulation: the case of ADHD', *Journal of Health Economics*, vol. 25(6), pp. 1094–118.
Dahl, G. and Lochner, L. (2012). 'The impact of family income on child achievement: evidence from the earned income tax credit', *American Economic Review*, vol. 102(5), pp. 1927–56.
Deshpande, M. (2016a). 'Does welfare inhibit success? The long-term effects of removing low-income youth from the disability rolls', *American Economic Review*, vol. 106(11), pp. 3300–30.
Deshpande, M. (2016b). 'The effect of disability payments on household earnings and income: evidence from the SSI children's program', *Review of Economics and Statistics*, vol. 98(4), pp. 638–54.
Dhuey, E., Figlio, D., Karbownik, K. and Roth, J. (2019). 'School starting age and cognitive development', *Journal of Policy Analysis and Management*, vol. 38(3), pp. 538–70.
Epple, D. and Romano, R. (2011). 'Peer effects in education: a survey of the theory and evidence', in (J. Benhabib, A. Bisin and M. Jackson, eds.), *Handbook of Social Economics*, vol. 1, pp. 1053–163, Elsevier.
Figlio, D., Guryan, J., Karbownik, K. and Roth, J. (2013). 'The effects of poor neonatal health on children's cognitive development', NBER Working Paper 18846.
Figlio, D., Guryan, J., Karbownik, K. and Roth, J. (2014). 'The effects of poor neonatal health on children's cognitive development', *American Economic Review*, vol. 104(12), pp. 3921–55.
Fletcher, J., Hair, N. and Wolfe, B. (2012). 'Am I my brother's keeper? Sibling spillover effects: the case of developmental disabilities and externalizing behavior', NBER Working Paper 18279.
Fletcher, J. and Wolfe, B. (2008). 'Child mental health and human capital accumulation: the case of ADHD revisited', *Journal of Health Economics*, vol. 27(3), pp. 794–800.
Guryan, J., Hurst, E. and Kearney, M. (2008). 'Parental education and parental time with children', *Journal of Economic Perspectives*, vol. 22(3), pp. 23–46.
Heckman, J. (2006). 'Skill formation and the economics of investing in disadvantaged children', *Science*, vol. 30(5782), pp. 1900–2.
Heissel, J. (2017). 'Teenage motherhood and sibling outcomes', *American Economic Review*, vol. 107(5), pp. 633–7.
Joensen, J.S. and Nielsen, H.S. (2018). 'Spillovers in education choice', *Journal of Public Economics*, vol. 157, pp. 158–83.
Kalil, A., Ryan, R. and Corey, M. (2012). 'Diverging destinies: maternal education and the developmental gradient in time with children', *Demography*, vol. 49(4), pp. 1361–83.
Karbownik, K. and Özek, U. (2019). 'Setting and good example? Examining sibling spillovers in educational achievement using regression discontinuity design', NBER Working Paper 26411.
Kimmel, J. (1998). 'Child care costs as a barrier to employment for single and married mothers', *Review of Economics and Statistics*, vol. 80(2), pp. 287–99.
Kreiner, C. and Sivertsen, H. (2018). 'Neonatal health of parents and cognitive development of children', CEBI Working Paper.
Krueger, A. and Schanzenbach, D. (2001). 'The effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from Project STAR', Economic Journal, vol. 111(468), pp. 1–28.
Kvist, A., Skyt Nielsen, H. and Simonsen, M. (2013). 'The importance of children's ADHD for parents' relationship stability and labor supply', *Social Science and Medicine*, vol. 88, pp. 30–8.
Landersø, R., Skyt Nielsen, H. and Simonsen, M. (2020). 'Effects of school starting age on the family', *Journal of Human Resources*, forthcoming.
Manski, C.F. (1993). 'Identification of endogenous social effects: the reflection problem', *Review of Economic Studies*, vol. 60(3), pp. 531–42.
Manski, C.F. (2000). 'Economic analysis of social interactions', *Journal of Economic Perspectives*, vol. 14(3), pp. 115–36.
Mazumder, B. (2008). 'Sibling similarities and economic inequality in the U.S.', *Journal of Population Economics*, vol. 21(3), pp. 685–701.
Mitchell, W. (2007). 'Research review: the role of grandparents in intergenerational support for families with disabled children: a review of the literature', *Child and Family Social Work*, vol. 12(1), pp. 94–101.
Moffitt, R.A. (1999). 'Models of treatment effects when responses are heterogeneous', *Proceedings of the National Academy of Sciences*, vol. 96(12), pp. 6575–6.
NCES. (2016). 'Digest of education statistics, 2014', National Center for Education Statistics 2016-006.
Nicoletti, C. and Rabe, B. (2019). 'Sibling spillover effects in school achievement', *Journal of Applied Econometrics*, vol. 34(4), pp. 482–501.
Noonan, K., Reichman, N. and Corman, H. (2005). 'Mothers' labor supply in fragile families: the role of child health', *Eastern Economics Journal*, vol. 31(4), pp. 601–16.

Ozier, O. (2018). 'Exploiting externalities to estimate the long-term effects of early childhood deworming', *American Economic Journal: Applied Economics*, vol. 10(3), pp. 235–62.
Parman, J. (2015). 'Childhood health and sibling outcomes: nurture reinforcing nature during the 1918 influenza pandemic', *Explorations in Economic History*, vol. 58, pp. 22–43.
Powers, E. (2003). 'Children's health and maternal work activity: estimates under alternative disability definitions', *Journal of Human Resources*, vol. 38(3), pp. 522–56.
Price, J. (2008). 'Parent–child quality time: does birth order matter?', *Journal of Human Resources*, vol. 43(1), pp. 240–65.
Qureshi, J. (2018a). 'Additional returns to investing in girls' education: impact on younger sibling human capital', Economic Journal, vol. 128(616), pp. 3285–319.
Qureshi, J. (2018b). 'Siblings, teachers and spillovers on academic achievement', *Journal of Human Resources*, vol. 53(1), pp. 272–97.
Raaum, O., Salvanes, K.G. and Sørensen, E. (2006). 'The neighbourhood is not what it used to be', Economic Journal, vol. 116(508), pp. 200–22.
Rege, M., Telle, K. and Votruba, M. (2011). 'Parental job loss and children's school performance', *Review of Economic Studies*, vol. 78(4), pp. 1462–89.
Reichman, N., Corman, H. and Noonan, K. (2004). 'Effects of child health on parents' relationship status', *Demography*, vol. 41(3), pp. 569–84.
Salkever, D. (1982). 'Children's health problems and maternal work status', *Journal of Human Resources*, vol. 17(1), pp. 94–109.
Seltzer, M., Greenberg, J., Floyd, F., Pettee, Y. and Hong, J. (2001). 'Life course impacts of parenting a child with a disability', *American Journal on Mental Retardation*, vol. 106(3), pp. 265–86.
Solon, G. (1999). 'Intergenerational mobility in the labor market', in (O. Ashenfelter and D. Card, eds.), pp. 1761–800, *Handbook of Labor Economics*, vol. 3A, Elsevier.
Stabile, M. and Allin, S. (2012). 'The economic costs of childhood disability', *Future of Children*, vol. 22(1), pp. 65–96.
Vinopal, K. and Gershenson, S. (2017). 'Re-conceptualizing gaps by socioeconomic status in parental time with children', *Social Indicators Research*, vol. 133(2), pp. 623–43.
Yi, J., Heckman, J., Zhang, J. and Conti, G. (2015). 'Early health shocks, intra-household resource allocation and child outcomes', Economic Journal, vol. 125, pp. F347–71.