IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11141

Spillovers in Education Choice

Juanna Schrøter Joensen
Helena Skyt Nielsen

NOVEMBER 2017

DISCUSSION PAPER SERIES

IZA DP No. 11141

Spillovers in Education Choice

Juanna Schrøter Joensen
University of Chicago and IZA

Helena Skyt Nielsen
Aarhus University and IZA

NOVEMBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Spillovers in Education Choice*

This paper examines how skills are shaped by social interactions in families. We show that older siblings causally affect younger sibling's education choices and early career earnings. We focus on critical course choices in high school and overcome the identification challenges of estimating spillover effects in education by exploiting exogenous variation in choice sets stemming from a pilot program. The pilot induced an essentially random subset of older siblings to choose advanced math-science at a lower cost, while not directly affecting the course choices of younger siblings. We find that younger siblings are 2-3 percentage points more likely to choose math-science if their older sibling unexpectedly could choose math-science at a lower cost. We argue that the main influence of the pilot program on the younger siblings may be attributed to the social influence of the older sibling. Spillovers are strongest among closely spaced siblings, in particular brothers, and they have a lasting impact on the career out-comes of younger brothers. We argue that competition is likely one of the driving forces behind younger siblings conforming to their older siblings' choices.

JEL Classification: I21, I24, J24

Keywords: social interaction, siblings, high school curriculum, skill formation

Corresponding author:
Juanna Schrøter Joensen
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
USA

E-mail: JJoensen@uchicago.edu

* Thanks to the editor and two anonymous referees for constructive comments that sharpened the pa-per. We also appreciate enlightening discussions with and comments from Anders Björklund, Tore El-lingsen, Ina Ganguli, Leonie Gerhards, Jim Heckman, Maria Knoth Humlum, Eric Maurin, Magne Mogstad, Julia Nafziger, Torsten Persson, Barbara Petrongolo, Alessandra Voena, and participants at ESWC 2015, SED 2015, SOLE 2014, ELE 2014, IWAEE 2013 and seminar participants at the CEP-LSE Educa-tion Group, Copenhagen Business School, Federal Reserve Bank of New York, IIES Stockholm Universi-ty, Lancaster University, Oslo University, Stavanger University, SOFI Stockholm University, Sussex Uni-versity, University of Chicago, University of Southern Denmark, York University, and Aarhus University. This paper was previously circulated under the title "Peer Effects in Math and Science". We thank Louise Broman, Kasper Jørgensen and Pernille Hansen for research assistance. The usual disclaimers apply.

1. Introduction

Social interactions may play an important role in the formation of skills. Social groups may either transmit information about particular educational investments or carry social norms and identity concerns influencing an individual's educational decision. Social interactions may reinforce or counteract the direct effect of economic shocks or policy interventions. In this paper, we focus on social interactions within the family in the context of advanced math-science choices in high school.[1] To the best of our knowledge, this is the first paper to quantify sibling spillovers of education choices *and* to document a considerable longer-term career impact.

Estimating causal spillover effects is challenging due to simultaneity, correlated unobservables, and endogenous peer group membership (Manski, 1993; 1995). We study naturally occurring social groups and exploit exogenous variation in the cost of taking up advanced math and science in high school among a partial population (Moffitt, 2001).[2] Advanced math and science courses are critical career choices that increase college access and improve future earnings prospects. Understanding what determines these choices is crucial for understanding income inequality, since prior research suggests that more than 50% of individual income variation over the life-cycle is explained by choices made prior to age 18 (Keane and Wolpin, 1997; Cunha et al., 2005; Huggett et al., 2011).

We exploit the fact that some older siblings in Denmark in 1984-87 were unexpectedly exposed to a pilot scheme after entering high school and investigate whether they influenced the course choices of their younger siblings. In previous work (Joensen and Nielsen, 2009; 2016) we investigated the direct impact of the lower cost of choosing advanced math and science on the individuals' education choices, subsequent careers, and earnings. Those unexpectedly exposed to the pilot scheme obtained more advanced college degrees and substantially higher earnings. In this paper, we analyze whether there are spillover effects on younger siblings who were unexposed themselves. We find that younger siblings are 2-3 percentage points more likely to choose math and science if their older sibling was exposed to the pilot scheme. If we invoke the exclusion restriction that the pilot scheme influences the younger sibling's course choice only through the older sibling's course choice, the influence of the pilot scheme on younger siblings' course choices can be interpreted as a causal peer effect. Since the

[1] This choice is a prerequisite for increasing the supply of college graduates in science, technology, engineering, and mathematics (STEM). Any policy aiming to increase the investment in such skills (e.g. increased course requirements implemented in the US following *A Nation at Risk*, Gardner et al., 1983) may be seriously dampened or amplified by social interaction effects. Social interaction effects are extremely important during the teenage years when decisions on more advanced coursework are taken (Akerlof, 1997; Akerlof and Kranton, 2002; Card and Giuliano, 2013).

[2] Our study is thus methodologically similar to the study of social interaction effects in program participation by Avvisati, Gurgand, Guyon and Maurin (2014), Dahl, Kostøl and Mogstad (2014), and Dahl, Løken and Mogstad (2014).

first-stage estimate is around 7 percentage points, this implies a peer influence of older siblings on younger siblings of about 0.3-0.5. Thus by affecting the choice of the older sibling, almost half of the direct effect is expected to spill over on the younger sibling's choice. We provide evidence that the bulk of the spillover of the pilot program runs through sibling interactions rather than other social interactions for closely spaced peers and other high school or social group characteristics. More generally, this suggests that knowledge about the social peer group is important to predict the total impact of education policies. Policies targeted at influential peers (such as older siblings) are amplified and have more widespread long-term effects in the presence of positive peer effects.

Our results suggest that there is substantial heterogeneity in peer effects – both in terms of how strongly the older sibling responds to incentives and how their choice spills over on the younger sibling. Peer effects are largest for relatively closely spaced siblings with up to four year age difference, in particular brothers. The longer-term effects on completed education and early career earnings are also large for younger brothers, but always insignificant for younger sisters. First-born siblings are the most influential peers and parental education is also important for spillovers. We provide suggestive evidence that sibling competition is likely driving the peer effect as younger siblings are less likely to conform to their older sibling's course choice if the older sibling is among the top performers.

Our paper speaks to the literature trying to unravel the role of the family in skill formation and human development more generally.[3] However, traditionally the child is not modelled as interacting with siblings, and rigorous economics research on the importance of social interactions among siblings is scarce.[4] We document that sibling spillovers in education choices are possibly large. Our results have implications for understanding equality of opportunity, inequality, and intergenerational mobility where the importance of family background for educational investments has long been recognized and sibling correlations have recently been examined.[5] We go beyond correlations and study sibling spillover effects in order to understand the complex family component that siblings share.

[3] See e.g. Becker and Tomes (1979, 1986), Cunha and Heckman (2007), Cunha et al. (2010) for a theoretical framework. Heckman and Mosso (2014) provide a recent review.

[4] Butcher and Case (1994) find that the presence of a sister in the sibship reduces education of daughters, and they argue that this is because a sister changes the reference group. Qureshi (2016) finds that the education of older sisters improves the education of younger brothers, and she argues that this result reflects improved quality of child care as the older sister takes care of younger siblings. Sibling spillover effects have also been documented from parental leave taking among brothers (Dahl, Løken and Mogstad, 2014), military service among brothers (Bingley, Lundborg and Lyk-Jensen, 2017), from child health (Black et al., 2017; Breining, Daysal, Simonsen and Trandafir, 2015) and adolescent smoking, drinking and marijuana use (Altonji, Cattan and Ware, 2017).

[5] See e.g. Black and Devereux (2010) and Solon (1999) for reviews on intergenerational correlations and Björklund and Jäntti (2012), Björklund and Salvanes (2010) and Mazumder (2008) for sibling correlations in schooling.

The remainder of the paper unfolds as follows: Section 2 discusses identification of social interaction effects and presents the institutional background which our empirical strategy relies on. Section 3 describes the data, while section 4 presents the empirical analysis of social interaction effects in the choice of math and science in high school. Section 5 investigates mechanisms and heterogeneity in peer effects. Section 6 concludes the paper.

2. Identification of Spillover Effects Using a High School Pilot Scheme

We exploit some unique features and changes in institutions in Denmark to identify sibling spillover effects. This section describes our identification strategy and the educational environment of the Danish high school. In the first subsection, we briefly explain the empirical challenge of identifying spillover effects and how we exploit the unique institutional setup to identify social interaction effects from older to younger siblings. Then we describe the two relevant high school regimes, which form the basis for our identification strategy. The second and third subsections concern the high school regime and the pilot scheme that provides us with exogenous variation in the cost of acquiring advanced math and science courses for the older siblings. The fourth subsection concerns the high school regime forming the basis for the math and science choices of their younger siblings.

2.1. Identifying Spillover Effects

Peer (or social interaction) effects occur when the choice of one individual affects the choices of other individuals in the same peer (or social) group. In this paper, we are interested in how math and science choices of an older sibling affect whether his or her younger sibling pursues advanced math and science courses. The general challenge of identifying peer effects lies in the empirical issues of: (i) endogenous group membership, (ii) simultaneity (the reflection problem), and (iii) correlated unobservables in the peer group.[6] These identification issues can be illustrated in a setting with only two individuals in each peer group - an older sibling and a younger sibling.

$$MathScience_{old} = \pi_0 + \pi_1 MathScience_{young} + \pi_2 X_{old} + \pi_3 X_{young} + \pi_4 F + \pi_5 S + \varepsilon_{old} \quad (1)$$

$$MathScience_{young} = \beta_0 + \beta_1 MathScience_{old} + \beta_2 X_{old} + \beta_3 X_{young} + \beta_4 F + \beta_5 S + \varepsilon_{young} \quad (2)$$

where the subscript $i \in \{old, young\}$ refers to sibling i in pair *f*.[7] $MathScience_i$ denotes whether sibling i chose advanced math and advanced science in high school, X_i denotes observable characteristics of sibling i, F denotes sibling pair specific characteristics like family background,

[6] Manski (1993; 1995) provides a more complete analysis of the identification of peer effects (or more generally endogenous peer effects), while Moffitt (2001) introduces the conceptual framework we adopt here.

[7] For ease of exposition, we suppress sibling pair subscript *f*.

gender composition, and age difference, whereas S denotes high school specific characteristics including quality and average family background. Finally, ε_i denotes other unobserved factors affecting the $MathScience$ choice of sibling i in pair f.

Our objective is to estimate a causal effect of the older sibling's $MathScience_{old}$ choice on the younger sibling's $MathScience_{young}$ choice. To be able to give a causal interpretation of the parameter estimate of β_1 in (2) we need to address the empirical issues (i)-(iii) mentioned above. The third issue of correlated unobservables is naturally a big concern in our setting, since siblings share many common social and genetic influences; including common genes, family background, neighborhood, and schools. All these common influences shape both siblings' preferences and abilities and could lead them to make similar high school course choices. Omitted variables bias due to contextual effects arises if we are not able to observe *all* the relevant sibling pair specific and individual variables. The first and the second issues are presumably minor in our setting: (i) siblings are born into the same family thus do not choose each other based on each other's characteristics and choices, and (ii) given the timing of high school course choices it seems plausible that the older sibling's course choice is independent of the younger sibling's choice ($\pi_1 = 0$) since the older sibling makes this choice years before the younger sibling. This exclusion restriction overcomes the reflection problem, as we postulate that the direction of the sibling effect goes from the older sibling to the younger sibling.[8] Nevertheless, this is not a necessary exclusion restriction as our empirical strategy addresses all these three empirical concerns, if the exogenous variation in the cost of acquiring advanced math and science for the older sibling is independent of both sibling pair specific factors, individual sibling characteristics, and unobserved social and genetic influences.

More specifically, our identification strategy exploits exogenous variation in the cost of acquiring advanced math and science stemming from a pilot scheme, where some older siblings unexpectedly got the option of a more flexible course combination. Let $PilotIntro_{old} = 0$ for older siblings in a traditional high school, where advanced math and science could *only* be achieved in a package of advanced math, advanced physics and intermediate chemistry. Let $PilotIntro_{old} = 1$ for older siblings in a pilot high school, where advanced math and science could *also* be achieved in a package of advanced math, advanced chemistry, and intermediate physics. This additional course package option

[8] The developmental psychology literature supports that the direction of behavioral influence goes from the older sibling to the younger sibling (Buhrmester, 1992). Bingley et al. (2017) test and reject that younger brother draft lottery affects older brother military service status. Altonji et al. (2013) also corroborate this assumption and impose it as an identifying assumption to estimate sibling influences on adolescence substance use.

was introduced unexpectedly just before the older sibling made the choice of advanced high school courses. Substituting into (1) and (2) we get:

$$MathScience_{old} = \pi_0 + \gamma PilotIntro_{old} + \pi_2 X_{old} + \pi_3 X_{young} + \pi_4 F + \pi_5 S + \varepsilon_{old} \quad (3)$$

$$MathScience_{young} = \beta_0 + \beta_1 MathScience_{old} + \beta_2 X_{old} + \beta_3 X_{young} + \beta_4 F + \beta_5 S + \varepsilon_{young} \quad (4)$$

The 2SLS estimate of β_1 in (4) can be interpreted as a causal sibling spillover effect. The key identifying assumptions are the exclusion restriction that $PilotIntro_{old}$ does not directly influence the younger sibling, i.e. is excluded from equation (4) *and* the orthogonality condition that $PilotIntro_{old}$ is independent of any individual (X_i) and sibling pair specific (F) or high school specific (S) variables, as well as any unobserved social and genetic influences on sibling choice (ε_i) for $i \in \{old, young\}$.[9]

We estimate γ as the first-stage regression coefficient in equation (3). This is the direct effect of the unexpected pilot introduction on the older sibling's *MathScience* choice.

We also examine whether the quasi-random variation in cost of *MathScience* for the older sibling changes the course choice of the younger sibling by estimating the following reduced-form:

$$MathScience_{young} = \delta_0 + \delta_1 PilotIntro_{old} + \delta_2 X_{old} + \delta_3 X_{young} + \delta_4 F + \delta_5 S + \varepsilon_{young} \quad (5)$$

where δ_1 can be interpreted as an "intention-to-treat" (ITT) effect of the unexpected pilot introduction (for the older sibling) on the younger sibling's *MathScience* choice. An advantage of the reduced form (5) is that it requires fewer assumptions to estimate the sign of the sibling spillover effect. If the pilot option is as good as randomly assigned such that $PilotIntro_{old}$ is uncorrelated with individual and sibling pair characteristics, the reduced form consistently estimates the causal effect of having an older sibling who unexpectedly could choose *MathScience* at a lower cost. To consistently estimate the size of the peer effect via 2SLS, we also need to assume that the *only* channel for younger siblings to be affected is through the older sibling's *MathScience* choice. In addition to this exclusion restriction, 2SLS when β_1 in (4) is a random coefficient also requires the monotonicity (or uniformity) condition that the pilot did not cause any older siblings to be less likely to choose *MathScience*. We are confident that the monotonicity assumption is reasonable in our application, since all options available at non-pilot schools were also available at schools introducing the pilot scheme.

Under the assumptions stated above, IV identifies the Local Average Treatment Effect (LATE) which is the causal effect of older sibling math-science choice on younger sibling math-science choice

[9] Moffitt (2001) labels this type of identification strategy a partial-population policy intervention. If $PilotIntro_{old}$ is randomly assigned, it will be uncorrelated with all observed and unobserved variables affecting each sibling's outcome in (3) and (4). This implies that γ can be identified from a regression of $MathScience_{old}$ on $PilotIntro_{old}$. As the younger sibling makes their choice after the older sibling, it also means that a consistent estimate of the spillover effect β_1 can be obtained by regressing $MathScience_{young}$ on $PilotIntro_{old}$ and scaling by γ (provided a strong instrument such that $\gamma \neq 0$). The orthogonality condition ensures that correlated unobservables do not bias the estimates.

for those older siblings who comply with the instrument; i.e. those who choose math-science when they are unexpectedly exposed to the pilot scheme and can do so at a lower cost, but would not have chosen it at the higher cost without the pilot scheme.

The orthogonality condition is corroborated in Joensen and Nielsen (2009; 2016) showing that $PilotIntro_{old}$ is independent of predetermined individual, family, and school characteristics for the students entering high school in 1984-87. This implies that older siblings are as good as randomly assigned to high schools which unexpectedly introduce the pilot scheme when they are enrolled in their first year and before they choose advanced courses for their second and third high school years. The following (sub)sections are devoted to provide further support for the key identifying assumptions. Subsection 2.3 shows that the instrument has a strong influence on the choice of math-science courses for cohorts entering high school in 1984-87 *and* that the pilot scheme was *not* implemented in schools with a student body that was more prone to choose math-science. We return to these empirical issues and additional empirical evidence to support the exclusion restriction which is unique to this paper in Section 4 and 4.2. in particular.

The following subsections describe the educational environment of the two relevant high school regimes: The Pre-1988 High School with restrictive course packages that the older siblings attended and the Post-1988 High School with much more flexible course choices for their younger siblings.

2.2. The Pre-1988 High School

In the period 1961-1988, the Danish high school system was a "branch-based" high school regime in which courses were bundled into restrictive course packages.[10] We focus on the cohorts entering high school in 1984-87. The main reason to focus on this period is that the supply of course packages provides us with relevant exogenous variation in the cost of acquiring advanced math and science for the older siblings. The combination of advanced math and science provides immediate access to STEM field college programs and it is by far the most lucrative course package; see Appendix A, Figures A1-A3. The variation in earnings across course packages is more important than the variation across high schools and more than doubles the high school premium. The decision about which package to opt for is taken at the end of the first year in high school. The only way to obtain advanced math and science was the package consisting of advanced math, advanced physics and intermediate chemistry, unless the student was enrolled at a pilot school, where the package could be adjusted to include advanced

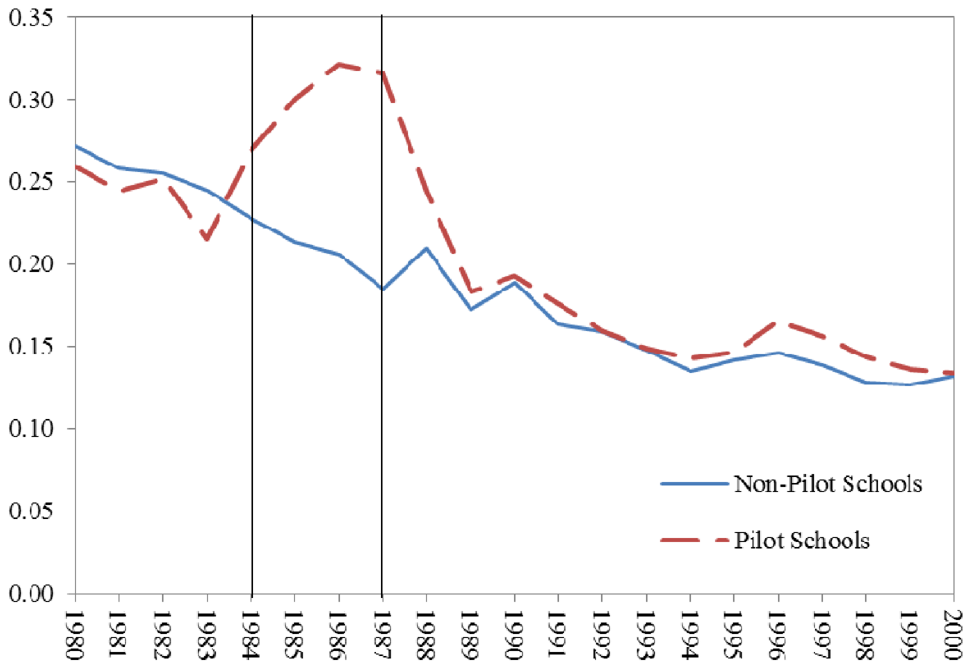
[10] Available course packages were labelled: Social Science and Languages, Music and Languages, Modern Languages, Classical Languages, Math-Social Science, Math-Natural Science, Math-Music, Math-Physics, and Math-Chemistry. The additional course package introduced at pilot schools was the Math-Chemistry option. See Joensen and Nielsen (2009) for additional details on choices sets and course content.

chemistry and intermediate physics instead. It is exactly this increased course flexibility which some students were unexpectedly exposed to that constitutes the quasi-experiment we exploit in this paper.

2.3. The Pilot Scheme

The pilot scheme was implemented as an experimental curriculum at about half of the high schools prior to the 1988-reform. The purpose of the pilot scheme was to test the impact of increased flexibility prior to the 1988-reform. Figure 1 illustrates the consequences of the pilot scheme on the course packages of the high school youth. Prior to the pilot scheme, the fraction choosing advanced math and science declined and was below 25% in 1983. The pilot scheme counteracted this declining trend by attracting youth to the alternative course package with a higher weight on chemistry and a lower weight on physics.

Figure 1. Fraction of High School Cohorts Choosing Math-Science, by School Pilot Status



Note: This Figure displays the fraction of each entering high school cohort in 1980-2000 who chooses advanced Math and Science. The blue solid line refers to the traditional high schools. The red dashed line refers to pilot schools, where pilot schools include all schools with pilot status at any point in time during 1984-87; i.e. 64 schools in total.

Table 1 gives an overview of the gradual implementation of the pilot scheme from 1984-87. The table is divided by types of high schools: schools with no pilot scheme (*PilotSchool=0*), schools where the pilot scheme was introduced after enrollment of the relevant cohort (*PilotSchool=0,* *PilotIntro=1*), and schools where the pilot scheme was implemented prior to enrollment of the relevant cohort (*PilotSchool=1, PilotIntro=0*).

Table 1. Introduction of the Pilot Scheme

High School	*Pilot School* = 0			*Pilot School* = 0 *Pilot Intro* = 1			*Pilot School* = 1 *Pilot Intro* = 0			All		
Cohort	N	*MathScience*	Schools	N	*MathScience*	Schools	N	*MathScience*	Schools	N	*MathScience*	Schools
1984	12,758	0.2343	122	3,145	0.3253	22	0	0.0000	0	15,903	0.2523	144
1985	10,645	0.2280	107	1,802	0.3330	15	3,069	0.3271	22	15,516	0.2598	144
1986	8,997	0.2120	91	1,749	0.3116	15	4,969	0.3490	37	15,715	0.2664	143
1987	8,297	0.1899	78	1,505	0.2811	12	7,553	0.3311	52	17,355	0.2593	142
Total	40,697	0.2187		8,201	0.3159		15,591	0.3360		64,489	0.2594	

Note: The table illustrates the introduction of the pilot scheme. For each of the affected cohorts, 1984-86, the table displays the number of schools which are traditional high schools only offering advanced Math with advanced Physics (*PilotSchool*=0), which are unexpectedly introducing the pilot scheme combining advanced Math with advanced Chemistry (*PilotSchool*=0, *PilotIntro*=1) and which already have adopted the pilot scheme (*PilotSchool*=1, *PilotIntro*=0).

Schools were not randomly assigned to become pilot schools, but we provide evidence that assignment was as good as random. From 1984-86, schools could apply to the Ministry of Education for permission to adopt the experimental curriculum and the ministry decided which shools were to adopt it, whereas in 1987 the high school principals could make this decision without approval from the ministry.[11]

It is clear, however, that students with a particular preference for chemistry may self-select into schools that are known to offer the pilot program before entrance. This is why we distinguish between students at pilot schools where the pilot scheme was unexpectedly introduced after they had enrolled (*PilotSchool=0*, *PilotIntro=1*) and those who knew that the school was a pilot school before they applied for admission to the school (*PilotSchool=1*, *PilotIntro=0*).

The instrumental variable strategy exploits the fact that the pilot scheme reduces the psychological cost of choosing advanced math and science since the students exposed to the scheme are free to choose either advanced physics and intermediate chemistry or advanced chemistry and intermediate physics.[12] Hence, first-year high school students enrolled at a school when it decided to introduce the pilot scheme were exposed to an unexpected exogenous cost shock, which induced more students to choose advanced math and science compared to students at non-pilot schools. If the selection of newly

[11] The schools which introduced the program in 1987 tend to be slightly negatively selected in terms of the students' math abilities, while no similar concerns are raised regarding the other cohorts. However, to maintain a large number of sibling pairs, we include the 1987 cohort of older siblings in the study, while checking the sensitivity of our results to leaving out this cohort. Similarly, the post-reform cohort enrolling pilot schools in 1988 appears to be slightly positively selected according to Figure 1. We thus also check the sensitivity of our results to leaving out this cohort.

[12] Traditionally, the opportunity cost of attending high school is interpreted as forgone earnings from unskilled work. We use a broader interpretation associated with time allocation across courses as well as between studies, leisure, and unskilled work. If students choose course combinations optimally given their preferences and abilities, then a more flexible choice set reduces the cost of taking a given course as there is a higher probability of a good match between feasible course combinations and the students' preferences and abilities.

participating schools is uncorrelated with the unobservables affecting math-science course choices, then the pilot scheme provides exogenous variation in students' math and science skills without influencing the outcome(s) of interest except through the effect on math and science choices.

The instrumental variable, *PilotIntro*, is equal to one if the individual enrolled in a high school which introduces the experimental curriculum for the first time when they were in their first year and about to choose advanced course packages for their second and third year, and it takes the value zero otherwise.

The instrument is strong if the unexpected introduction of the pilot scheme induces students to choose advanced math and science. This can be seen from Figure 1 and Table 1, and is directly tested and validated in Section 4.

The instrument satisfies the orthogonality condition if the pilot scheme is randomly assigned to schools and if individuals are randomly distributed across schools that have not yet decided to introduce the experimental curriculum. We minimize issues of students self-selecting into schools by focusing on those who were already enrolled, but had not yet chosen their final math-science level when the pilot scheme was unexpectedly introduced. The orthogonality condition is also violated (a) if the schools participate in the program based on the math abilities of local students or (b) if schools change as a consequence of the scheme: if the school develops an expertise in science or if the quality of teachers changes as a consequence of the pilot scheme.

Table 2. Placebo Tests on Pre-Pilot Cohorts

	Parameter Estimates (Standard Errors)		
	(1)	**(2)**	**(3)**
	First-stage: *MathScience*		
Placebo *PilotIntro*	-0.016 **	0.002	-0.005
	(0.006)	(0.005)	(0.005)
PilotIntro (actual)	1984-1985	1984-1986	1984-1987
PilotIntro (placebo)	1982-1983	1981-1983	1980-1983
Pre-pilot placebo period	t-2	t-3	t-4
Number of Individuals	**35,870**	**54,018**	**70,771**

Note: Parameter estimates and (standard errors) of the placebo pilot scheme introduction are displayed for first-stage OLS regressions of *MathScience* choice. The columns differ in which cohorts the placebo pilot introduction is assumed for and which schools are included in the sample. Column (1) leads *PilotIntro* two cohorts by assuming it was implemented for those who are graduating from the relevant school in the year it was actually implemented. Columns (2) and (3) lead *PilotIntro* three and four cohorts, respectively, by assuming it was implemented for those who have already graduated from the relevant school in the year it was actually implemented. The full set of cohort and parental background control variables is included. Significance at a 1%-, 5%-level and 10%-level are indicated by ***, ** and *, respectively.

It is not possible to directly test whether the pilot schools represent a sample of schools which is essentially random with respect to potential students' math and science ability, but we corroborate that this is a reasonable assumption. Table 2 presents placebo tests that support instrument exogeneity by showing that (a) is not a core concern. The table presents first-stage estimates (falsely) assuming that the pilot scheme was introduced two, three, or four years, respectively, prior to when it was actually introduced. The first specification assumes that the pilot scheme was implemented for the cohorts who were in their third and final year when the pilot was actually adopted, while the last two specifications assume that it was implemented for recently graduated cohorts. Neither of these cohorts should be affected, since they should already have made their final course choices before the pilot scheme was adopted. We find a small significantly negative effect of the pilot schemes for the first placebo test. This suggests that the schools introducing the pilot in 1984-85 had slightly (1.6 percentage points) fewer students choosing math and science before the school adopted the pilot program. The coefficient is small and the picture is consistent with what is also seen in Figure 1.[13] However, in the two last specifications where the distance to the actual reform is longer, the effect is smaller and insignificant. We therefore conclude that there existed only minor, if any, systematic prior differences in choices at schools which adopted the pilot scheme, and if anything, they should work in the opposite direction of the pilot.[14] In Section 3 below, we further test for similarities of student and parent characteristics across the school status, and we find almost no significant differences in characteristics determined prior to high school.

Second, the exclusion restriction is violated if (b) schools change as a consequence of the pilot scheme. If the school develops an expertise in science or if the quality of teachers changes as a consequence of the pilot scheme. Such effects would presumably most strongly influence the younger siblings if they attend the same high school as their older siblings and could confound the sibling peer effects. We cannot completely rule out such an effect. However, Figure 1 shows that the pilot scheme was introduced following a declining trend in the fraction choosing advanced math and science, and therefore qualified teachers for these courses would most likely be available in pilot schools as well as non-pilot schools during the pilot period and afterwards. Furthermore, the relevant compliers would switch from intermediate chemistry and intermediate math (and e.g. advanced biology and geography, depending on their alternative choice) to advanced chemistry and advanced math (and e.g. intermediate

[13] We have also graphed the red dashed line in Figure 1 separately by schools who introduce the pilot in 1984, 85, 85, and 87 separately. This reveals that the schools implementing the pilot in 1984 are driving the dip, but this is the cohort carrying the lowest weight in our main estimation sample of relatively closely spaced siblings.

[14] We have also performed the placebo tests for older siblings and for older brothers and sisters separately. For these subsamples, none of the parameters of main interest are significant, but power is also lower.

biology and geography, depending on their alternative choice). This means that 5-6 weekly chemistry and math lectures replace biology and geography. From the point of view of the complying student, this means that 20% of their weekly time schedule would be different during the 2nd and 3rd year of high school. From the point of view of the school, this means that the teaching supply should be adjusted corresponding to approximately 1% of the weekly course supply.[15] In the empirical analysis in Section 4.2 we further investigate to what extent the spillover effects may be going through the high school or other social interactions - such as common friends across school cohorts - rather than through sibling interactions. Our findings suggest that the bulk (at least 70%) of the spillover effect goes through social interactions among siblings.

Our instrument exploits the exogenous variation in the exposure of students to the option of switching the levels of physics and chemistry. Hence, the *treatment* of the older sibling that we investigate is the combined treatment of advanced math with advanced chemistry and intermediate physics. We cannot separately identify the effect of math and science courses from the potential synergy between them.

2.4. The Post-1988 High School

In 1988 there was an extensive structural reform of the Danish High School, which was the most fundamental high school reform since 1903. The reform abolished the "branch based" regime and substituted it with a "choice based" regime, where the main distinction is between mathematical and linguistic track students. The reform implied an extended choice set in the form of more flexible opportunities to combine optional courses.[16] In particular, the mathematical track students have the option of combining advanced math with any other advanced course; for example physics, chemistry, biology, social science, or a language course. This is the regime within which the younger siblings in our sample make their education choices. We focus on the younger siblings' choice of advanced math with advanced physics and/or advanced chemistry, since these are comparable to the relevant course combinations for the older sibling attending high school in the pre-1988 regime.

[15] This adjustment corresponds to around 32 additional math and/or chemistry teachers in the entire country in year 1987 when the program was at the maximum and 64 schools had introduced the pilot program; see Table 1, and most of these should be available as the declining math-science trend continued at the traditional high schools. This was a relatively small adjustment of the teaching supply in comparison to the subsequent adjustment due to the major reform in 1988, which is described in the next subsection.

[16] The reform also implied more weight on the high schools' role of preparing students for college, more required readings, more written assignments, more stringent non-attendance regulation, more grading, and more hours of instruction allocated to the compulsory courses.

In the post-1988 high school regime, students choose either the mathematical or the linguistic track upon entry. Each course is either common to all students on the chosen track (compulsory courses), compulsory for some and optional for others, or exclusively optional. The optional courses can be obtained at either advanced or intermediate level reflecting the complexity of the content, the number of lessons per week and the intensity of exams (written and/or oral).

All students are required to follow at least two (and at most three) optional advanced courses, and for the mathematical students there was a minimum required amount of math-science content, while for the linguistic students there was a minimum required amount of language content. The first year of high school consists only of compulsory courses (common as well as track-specific courses) taught in classes of at most 28 students. The second and third year of high school added at least three and at most four optional courses.[17] In addition to the requirements of at least two advanced optional courses, there were some bonds between some courses in order to preserve the possibility for courses to complement each other.

We follow younger siblings in this high school regime until the entry cohort of 1997. We focus on the younger siblings' choice of advanced math with either advanced physics or advanced chemistry, since these are comparable to the relevant pre-1988 regime course combinations of the older siblings.[18] Thus $MathScience_{young}$ in equation (4) is an indicator for whether the younger sibling chooses to combine advanced math with either advanced physics or advanced chemistry.

3. Data Description

3.1. Sample Selection

For our empirical analysis we use a panel data set comprising the population of individuals enrolling in high school from 1980 and onwards. The data are gathered from administrative registers and administered by Statistics Denmark. The data include basic demographic information such as date of birth, place of residence, and gender. What is crucial for this study is that we observe which schools offered the pilot scheme when, and we can identify which school the individual attended as well as the

[17] The compulsory courses common to all students are advanced Danish and history, intermediate English and basic physical education, biology, geography, religion, music, (visual) art, and ancient history. Track-specific compulsory courses for mathematical students comprise intermediate math and physics, basic chemistry, and a second foreign language. For the linguistic students the track-specific compulsory courses are basic natural sciences (including math) and Latin, as well as two other foreign languages. Commonly available optional intermediate courses comprise: biology, geography, chemistry, technical science, business and economics, drama, sports, and movie science, while optional advanced courses include all feasible continuations of the intermediate courses.

[18] Some curriculum changes are introduced with the reform, e.g. a historical dimension was incorporated into the math course while some advances in the experimental direction were incorporated into the physics course.

chosen course package. Furthermore, we have information on the dates for entering and exiting a high school education, along with an indication of whether the individual completed the education successfully, dropped out, or is still enrolled as a student as well as subsequent career outcomes. We augment this data with background information about the parents; including educational achievement and gross income. This information is recorded when the individual was 15-years old, which is prior to enrolling in high school.

The sample consists of individuals who are directly influenced by the quasi-experimental variation due to the gradual introduction of the pilot scheme for cohorts entering high school 1984-1987. From this sample, we select high school graduates who finished in three years and have a younger sibling who entered high school after 1987 and finished in three years.[19] In our main analysis, we focus on a homogeneous sample of relatively closely spaced sibling pairs (cohorts 1988-91, age gap ≤ 4 years), and for comparison we also report results for more widely spaced sibling pairs (cohorts 1988-97, age gap ≤ 10 years).[20]

3.2. Outcome and Control Variables

The outcome of interest is whether the post-reform peers choose to combine advanced math and science or not. Table B3 reveals a strong correlation in the choice of this course package across siblings: 28 % (14 %) of younger siblings chose this course package when the older sibling did (did not) choose this package, and the correlation varies across gender composition of the sibship.

Table B4 shows variation in the choice of advanced math and science when we distinguish between whether the older sibling was exposed to the pilot scheme or not. The proportion of younger siblings who chose this course package is 18% when the older sibling was not exposed ($PilotSchool$=0, $PilotIntro$=0) and 22% when the older sibling was unexpectedly exposed to the pilot scheme ($PilotSchool$=0, $PilotIntro$=1). The relationship appears to be very strong among pairs of brothers.

The control variables are crucial, as they should include key variables reflecting common influences in the environment that shape both siblings' high school course choices (as discussed in section 2). We include entry cohort fixed effects, sibling pair gender composition, parental background, county fixed effects and high school specific controls. Importantly, county fixed effects and predetermined high

[19] About 40% of a birth cohort attended the academic high school track at this point in time; hereof 10% do not complete in three years. The main part of drop out takes place before the choice of advanced math and science course packages. For older as well as younger siblings, dropout is uncorrelated with pilot school status.

[20] An overview of the sample selection is given in Table B1 in Appendix B. An overview of the distribution of sibling pairs across the older siblings' exposure to the pilot scheme for each high school cohort of younger siblings is given in Table B2 in Appendix B.

school means are thought to proxy permanent regional or high school specific differences in the quality of science teachers or the expertise in science teaching. Parental background variables are thought to proxy family variation in the preferences or ability for high school courses. They include a set of mutually exclusive indicator variables for the level of highest completed education of the mother and father, respectively, and their income as observed at the end of the year before the individual started high school.[21] Since this study is specifically concerned with spillovers of math and science course choices, we also include indicators for whether the mother's and father's highest completed education is within a STEM field or not in order to approximate a family-specific ability or preference for STEM fields. We have constructed indicator variables based on a definition of STEM fields, which follows the definition by the National Science Foundation (NSF) and includes math, engineering, natural and technical sciences as well as some social sciences and life sciences.

Table B5 in Appendix B shows descriptive statistics of background variables across pilot school status. From this table it is evident that the students whose older siblings entered high schools which had already adopted the pilot program ($PilotSchool{=}1$, $PilotIntro{=}0$) prior to their entry are potentially non-randomly selected, while those whose older siblings were unexpectedly exposed to the program ($PilotSchool{=}0$, $PilotIntro{=}1$) are not systematically different from students at schools without the pilot program ($PilotSchool{=}0$, $PilotIntro{=}0$). This lends further support for the exogeneity of $PilotIntro$ as an instrument for older siblings' course package choice as there is no significant selection on pre-determined obervables.

4. Estimates of Sibling Spillover Effects

Table 3 presents the main results from the empirical analysis of how an increase in the incentives to do math-science for the older sibling spills over on the younger sibling's math-science choice.

As a benchmark, the OLS estimates of β_1 in equation (2) show a strong positive association between math-science choices of older and younger siblings. The probability of a younger sibling choosing math-science is 14 percentage points higher if the older sibling chose math-science. This is true for all three specifications. Column (1) only controls for whether schools had already implemented the pilot scheme. Column (2) adds sibling pair gender composition (which is the most influential control variable), cohort fixed effects, and parental background variables. Column (3) adds county fixed effects and predetermined high school means of parents' highest completed education and income, which are thought to proxy permanent high school specific effects such as the teacher quality.

[21] The variables mentioned so far are identical to the variables included in Joensen and Nielsen (2009; 2016).

Reduced form OLS estimates of δ in equation (5) are also presented in columns (1)-(3) in Table 3. The reduced form estimates suggest that there are spillover effects from the introduction of the pilot scheme on the younger siblings. The effects are statistically significant when the age difference between siblings is less than four years, but not when the distance is up to ten years. When the age difference is less than four years, the magnitude of the estimates ranges from 3.6 to 3.0 (when sibling pair gender composition, cohort fixed effects, and parental background variables are included) to 2.1 percentage points and insignificant (when county fixed effects and high school specific variables are added). Figure B1 in Appendix B illustrates point estimates and confidence intervals when the maximum age difference between siblings varies. The point estimates are smaller for widely spaced as well as closely spaced siblings, and it is literally zero for the few siblings less than 2 years apart in specification (3).

While the reduced form estimates are informative about spillover effects of the pilot program, the IV estimates are informative about spillover effects of the program going through the older siblings' course choice if we invoke the exclusion restriction that the only channel through which $PilotIntro$ affects $MathScience_{young}$ is through its effect on $MathScience_{old}$. We present first-stage estimates of γ in equation (3) and 2SLS estimates of β_1 in equation (4) in columns (4)-(6) in Table 3.[22] The first stage corroborates instrument relevance as older siblings who were unexpectedly exposed to the pilot increased their probability of choosing math-science by 7 percentage points. The IV estimates suggest a strong sibling spillover effect, since the younger siblings whose older siblings are incentivized to take math-science because of the pilot are 0.3-0.5 percentage points more likely to also choose math-science. Again, the point estimates are not significantly different across specifications, only statistically significant for closely spaced siblings, and not significantly different from zero once all controls are included.[23]

[22] The main results are robust to alternative estimation methods. We report the results from the (bivariate) probit estimator in Table B6 in Appendix B. Our main results are robust and the first-stages are literally unchanged. Conclusions from the bivariate probit model appear slightly stronger; particularly in the subgroup analysis in Section 5. However, as in Altonji, Elder and Taber (2005) identification is mainly driven by the parametric assumptions when covariates are included in the bivariate probit model. In addition, we have used the semi-parametric estimator by Abadie (2003) as a robustness check. The IV estimate without control variables is statistically indistinguishable from the bivariate probit (marginal effect 0.384 vs. 0.360) and 0.310 with all controls in specification (6).

[23] We also examine how the 2SLS estimates change when the maximum age difference between siblings varies. The first-stage point estimates do not change as maximum age difference increases from two to ten years. The 2SLS point estimates get smaller for widely spaced siblings and follow the same pattern as the reduced form estimates presented in Figure B1 in Appendix B, but are larger and noisier. This suggests that the older sibling's choice is largely independent of how age-distant the younger sibling is, but the spillover from older to younger sibling becomes weaker as the age distance increases beyond four years. Our estimates are too noisy to draw conclusive inference on these patterns.

Table 3. Estimates of Spillover Effects: Main Results

		Parameter Estimates (Standard Errors)					
		OLS			2SLS		
	N	(1)	(2)	(3)	(4)	(5)	(6)
		First-stage: Older Sibling *MathScience*					
PilotIntro	7,786				0.068 ***	0.069 ***	0.069 ***
Younger Sibling 1988-91, ≤4y					(0.018)	(0.017)	(0.017)
PilotIntro	17,691				0.067 ***	0.073 ***	0.074 ***
Younger Sibling 1988-97, ≤10y					(0.012)	(0.012)	(0.012)
		Reduced-form: Younger Sibling *MathScience*					
PilotIntro		0.036 **	0.030 **	0.021			
Younger Sibling 1988-91, ≤4y		(0.015)	(0.014)	(0.015)			
PilotIntro		0.010	0.009	0.006			
Younger Sibling 1988-97, ≤10y		(0.010)	(0.009)	(0.009)			
		Outcome: Younger Sibling *MathScience*					
Older Sibling *MathScience*		0.140 ***	0.143 ***	0.142 ***	0.523 **	0.436 **	0.298
Younger Sibling 1988-91, ≤4y		(0.010)	(0.010)	(0.010)	(0.237)	(0.215)	(0.213)
Older Sibling *MathScience*		0.139 ***	0.143 ***	0.143 ***	0.143	0.119	0.076
Younger Sibling 1988-97, ≤10y		(0.007)	(0.007)	(0.007)	(0.139)	(0.123)	(0.126)
Control Variables:							
Entry Cohort Fixed Effects			+	+		+	+
Sibling Pair Gender Composition			+	+		+	+
County Indicators				+			+
Parental Variables (for Mother and Father):							
Highest Completed Education and Income			+	+		+	+
Highest Completed Education in STEM Field			+	+		+	+
HS Mean of Highest Completed Education and Income				+			+

Note: Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

4.1. Counting and Characterizing Compliers

The IV point estimates are larger than the OLS point estimates, although not significantly so. At a first glance, this may be surprising as we suspected that common unobservables like shared genes and social environment may affect both older and younger siblings' math-science choices in the same direction. Thus implying that the OLS estimate is upward biased. However, it may suggest that there is substantial heterogeneity in causal peer effects and older siblings who are at the margin of choosing math science are more influential for their younger siblings than others. This would be consistent with sibling competition: if the older sibling is a "math-science star", the younger sibling would be more reluctant to compete than if the older sibling is on the margin of choosing math-science. We examine heterogeneity in more detail in Section 5, and provide some support for this claim.

Now we turn to counting and characterizing compliers and decomposing the OLS and IV estimates in order to better understand their differences and what the estimated spillover effects embody.

First, we count and characterize compliers. Let $MathScience_{young,1}$ be the potential math-science choice of the younger sibling if the older sibling chose math-science ($MathScience_{old}=1$) and let $MathScience_{young,0}$ be the potential math-science choice otherwise ($MathScience_{old}=0$). Likewise, let $MathScience_{old,1}$ be the potential math-science choice of the older sibling if unexpectedly getting the low cost math-science option ($PilotIntro=1$) and let $MathScience_{old,0}$ be the potential math-science choice otherwise ($PilotIntro=0$). Under the monotonicity assumption stated in Section 2, the potential outcome framework splits sibling pairs into three groups: Never-takers are older siblings who do not choose math-science even when exposed to the pilot program; i.e. $MathScience_{old,1}=0$, $MathScience_{old,0}=0$. Always-takers are older siblings who choose math-science even without the pilot program; i.e. $MathScience_{old,1}=1$, $MathScience_{old,0}=1$. Compliers are individuals who choose math-science when they are unexpectedly exposed to the pilot program, but not when they do not have the pilot option available; i.e. $MathScience_{old,1}=1$, $MathScience_{old,0}=0$. Table B2 (bottom) shows that 34.57% of 865 students enrolled in high schools unexpectedly offering the pilot program chose math-science, while only 27.72% of 4,463 students enrolled in high schools without the pilot program did. Under the monotonicity assumption, this leaves a total of 6.85% or 364 compliers, among whom 59 attended a high school unexpectedly offering the pilot program. Parents of compliers more often have a 4-year college degree or more education, mother's education is more often within STEM fields, and compliers are more often males and have higher high school GPA compared to the population of high school graduates; see Table B7 in Appendix B.

Second, we decompose the OLS and IV estimates. The OLS estimates simply compare the younger sibling math-science probability for always-takers and compliers at pilot schools to the younger sibling math-science probability for never-takers and compliers at non-pilot schools. This difference is always around its raw data moment presented in Table B3 in Appendix B. That is, the difference between $\mathrm{P}(MathScience_{young}=1 \mid MathScience_{old}=1) = 0.283$ and $\mathrm{P}(MathScience_{young}=1 \mid MathScience_{old}=0) = 0.144$. This difference is in line with the OLS point estimates of β_1 around 0.14 presented in columns (1)-(3) in Table 3.

With essential heterogeneity, IV identifies the LATE; i.e. the expected causal effect for compliers $E[\,MathScience_{young,1}-MathScience_{young,0} \mid MathScience_{old,1}=1,\ MathScience_{old,0}=0\,]$ and can be decomposed into the expected potential outcomes with and without treatment for the compliers, respectively; i.e. $P(\,MathScience_{young,1}=1 \mid MathScience_{old,1}=1,\ MathScience_{old,0}=0\,)$ and $P(MathScience_{young,0}=1 \mid MathScience_{old,1}=1,\ MathScience_{old,0}=0\,)$. We use the method suggested by Abadie (2002, 2003) to estimate these expected potential outcome of compliers *with* and *without* treatment to be 0.521 and 0.211, respectively, for our specification (6) in Table 3.[24]

If we compare these estimates of complier potential choice probabilities to the observed probabilities of younger siblings choosing math-science conditional on older sibling choice and pilot status presented in Tables B3 and B4, the expected probability of compliers *without* treatment (0.211) is 50% higher than the observed probability when the older sibling does not choose math-science $P(MathScience_{young}=1 \mid MathScience_{old}=0)=0.144$ and as high as this probability for brother-brother pairs (0.206) and younger siblings whose older siblings were unexpectedly exposed to the pilot; i.e. $P(MathScience_{young}=1 \mid PilotIntro=1)=0.216$. More strikingly, the expected probability of compliers *with* treatment (0.521) is almost twice as large as the observed probability when the older sibling chooses math-science $P(MathScience_{young}=1 \mid MathScience_{old}=1)=0.283$ and 10 percentage points larger than this probability for brother-brother pairs (0.423). This strongly suggests that the younger siblings of the complying older siblings were slightly more prone to choose math-science despite their older siblings not doing so, but way more prone to conform and follow in their older sibling's footsteps *if* the older sibling is induced to choose math-science because of the pilot program.

The IV estimates become so large because the instrument affects a subgroup for which the younger sibling is extremely likely to conform to the older sibling's choice *if* they choose math-science. In this respect, the OLS estimates place a higher weight on the younger siblings who are about half as likely to conform to the math-science choice of the always-taker older sibling (0.283 versus 0.521). This large difference between compliers and always-takers in expected potential outcomes *with* treatment is thus the main reason why the IV estimates are so much larger than the OLS estimates.

4.2. Potential Threats to Identification

We now turn to providing some additional empirical support of the exclusion restriction and exogeneity of $PilotIntro$, which is imperative for the causal inference based on the IV estimates.

[24] The decomposition renders similar conclusions if we instead focus on the other specifications (4) and (5), where the difference is even larger. These differences are also in line with the IV estimates of in columns (4)-(6) in Table 3.

We already presented evidence for instrument exogeneity in Section 2.3. First, showing that the pilot program was not introduced at schools where the student body was more prone to choose math-science (Table 2), if anything less prone. Second, we presented evidence showing that almost no predetermined characteristics differ significantly across schools (Table B5); two exceptions (of 32) are fathers being 2 percentage points more likely to have compulsory schooling only and students being 3 percentage points more likely to be in a large sibship of four or more siblings. Under the null hypothesis that there are no predetermined differences between schools which do and do not adopt the pilot, we should consider it a reasonable size of the test if we reject a true null hypothesis (and make a Type I error) 6.25% of times.

The exclusion restriction is arguably a much stronger assumption when analyzing sibling spillovers than when analyzing the direct effects on older siblings' subsequent education and labor market outcomes. In this context, it states that the *only* channel through which the pilot program affected the younger sibling's choice was through social interactions with the older sibling. We provide five pieces of empirical evidence to corroborate that the exclusion restriction is still reasonable in this context of sibling spillovers in math-science choice. The first two sensitivity checks aim at ruling out other direct social interactions while the last three pieces of evidence aim at ruling out the school itself as an important transmitter of the pilot spillover effect.

Those who are in adjacent high school cohorts or close in age could directly influence each other's choices as they may share the same social environment in terms of leisure activities or common friends. First, we therefore exclude older siblings enrolling high school in 1987 and younger siblings enrolling high school in 1988, respectively. Our results are robust to this exercise and, if anything, stronger in these subsamples; see Table B8 in Appendix B. Second, we exclude sibling pairs with an age difference of less than two years. The results are reported in Table B9 in Appendix B and, if anything, stronger for this subsample.

It is also possible that high schools (or counties) with or without traditions for math and science act as catalysts for sibling spillovers. Third, we therefore distinguish between high schools (and counties) with a low versus high fraction of students choosing math and science before the pilot was introduced; see Table B10 in Appendix B. Even though pilot schools on average have a slightly lower fraction of students choosing math and science prior to the pilot (see Figure 1), our results indicate that spillovers are indeed higher at high schools with a prior tradition for math and science. This pattern is weaker, but also present, when we instead distinguish between counties with or without a prior tradition for math and science. This is consistent with Section 4.1 as compliers in environments with a stronger tradition for math and science may face a stronger social pressure or a higher social benefit to conform to the modal choice in their social peer group at large.

It is also possible that the pilot changed school quality or that particular parental characteristics drive choices. Fourth, we therefore present results from a placebo test where we match only children from entry cohorts 1984-87 with only children from entry cohorts 1988-91; see Table B11 in Appendix B. In the left panel, we match the children randomly based on mother's education and income, and in the right panel, we match only children attending the *same* high school. We perform exact matching without replacement. In both cases the correlation between their course choices are positive but insignificant. When we match only children attending the same high schools, the reduced-form and the 2SLS estimates tend to be larger than when they are randomly matched across high schools. However, the estimates are insignificant and less than 30% of the main spillover estimates in Table 3. It should also be noted that the OLS estimates for the randomly matched siblings with similar mother characteristics and attending the same school are just over 30% of the OLS estimates for the real siblings in Table 3. We conclude that the potential spillover effect transmitted through the school is quantitatively much smaller than the one transmitted through siblings.[25]

Fifth, Figure B2 in Appendix B presents evidence that measures of high school quality did not change because of the pilot. Potential long-term impacts of the pilot program on school quality could also confound the estimated sibling spillovers as 81.5% of younger siblings attend the same high school as their older siblings. Panel (a) shows that the number of students graduating within three years – the ordained high school duration – fluctuates similarly by pilot status. The cohorts around the major high school reform (1987-88) may be exceptions to this rule, since the pilot schools tend to produce slightly more on-time graduates in these two years. However, as shown in Table B8, our main conclusions are robust to excluding these two cohorts. Panel (b) shows that log parental income follows a similar increasing trend, while panels (c)-(f) show that this is also true for the fraction of mothers and fathers with a longer than 4-year college degree and STEM degrees.

These five pieces of empirical evidence largely corroborate that the exclusion restriction is reasonable and that social interactions with the older sibling are by far the most important source of the spillover effect, while other social peer interactions and school quality concerns should only be minor transmitters.

[25] It should be noted that only children attending the same high school may know each other and socially interact through extra-curricular activities like sports clubs. This direct social interaction is less likely if their age difference is larger. The estimates are not significantly affected by leaving out placebo siblings within the same high school with an age difference of less than two years.

5. Understanding Heterogeneity in Sibling Interactions

In this section, we seek to better understand heterogeneity in pilot incentive responses and spillover effects. We examine which older siblings respond more (or less) to changes in incentives, as well as for which sibling pairs the spillover effect is stronger (or weaker). We explore differences in complier characteristics and spillover effects across gender, ability, sibship composition, and parental background. These heterogeneous effects provide suggestive evidence on mechanisms.

The student may gain utility from behaving similarly to an older sibling who inspires academic aspirations or compete with the older sibling in terms of academic achievements in order to improve his or her identity payoff. The social environment may also allow for information sharing in which case a student with an older sibling who pursued math and science faces less uncertainty about the difficulty and joy of this course package and about the future prospects compared to other students.

In order to shed light on which mechanism is more important, we draw upon psychological and sociological literature on the social interactions among siblings. This literature focuses on gender composition, birth spacing and birth order as fundamentally important characteristics explaining the nature of social interaction between siblings.

The importance of birth order was first mentioned by Adler (1927) and has been found for education and cognitive outcomes (Björklund and Jäntti, 2012; Black et al., 2005). Adams (1972) suggests that second and middle children often try to catch up with the first child and compete, but the youngest child less so. In Appendix C we scrutinize birth order effects and show that first borns exert a strong influence on younger siblings which seems to be driven by strong direct spillover on the second born.

Conley (2000) stressed that same-sex sib ships are more competitive and achievement-oriented than other sibships, and in particular if they consist of two boys. Grose (1991) states that two closely spaced brothers produce the most rivalry. Adams (1972) suggests that siblings who are less than five years apart are more competitive, while siblings who are more than five years apart tend to behave like separate sibships. Thus, it seems plausible that sibling rivalry and competition is a common denominator which may be particularly important among closely spaced pairs of brothers.

High school course choices may reflect competitive actions. Various characteristics of math, in particular, but to some extent also science, suggest that it is a competitive discipline (Niederle and Vesterlund, 2010). In the discipline of math, answers are either right or wrong, which makes it easier to claim victory. Furthermore, math skills are strong predictors of future performance, which means that the monetary gains from excellent performance may be sizeable.[26] The math discipline attracts

[26] See e.g. Altonji (1995), Cortes et al. (2015), Falch et al. (2014) and Joensen and Nielsen (2009; 2016).

more males who are known to be attracted to competition, while females tend to shy away from mixed-sex competition and to do worse in high-stake mixed-sex competition (Niederle and Vesterlund, 2007). Buser et al. (2014) estimate that around 20% of the gender difference in the choice of an academic math and science high school track is explained by gender differences in competitiveness. Therefore, *MathScience* course choice may be a battlefield characterized by competition and sibling rivalry. To investigate the importance of sibling rivalry, and other potential mechanisms, we study heterogeneity of effects across subgroups divided by sibship gender composition and ability.

5.1. Ability and Gender Composition of Sibship

Table 4 presents estimates separately by gender of the older and younger sibling, respectively. The first stage results show that older brothers are much more likely to comply with the instrument than older sisters, and older siblings in sibships with younger brothers are also more likely to comply with the instrument than those with younger sisters. The pilot scheme inducing adolescents to choose advanced math and science thus influenced students in sibships containing boys more than others. For example, the relative likelihood that a complier is an older brother with a younger brother is around 1.8 and given by the ratio of the first-stage coefficient on *PilotIntro* for the subsample of older brothers relative to the estimate for the full sample. This means that older brothers with younger brothers are around 80% more likely to comply with the unexpected introduction of the pilot scheme than the average older sibling. The reduced-form estimates reveal a very strong effect on younger brothers: when their older sibling was unexpectedly exposed to the pilot, younger brothers' probability of taking up advanced math and science increased by 6-9 percentage points. The sibling spillover effects also differ by gender: the point estimates are largest and only statistically significant for brother–brother pairs as younger brothers are 70 percentage points more likely to choose math-science if their older brother unexpectedly chose to do so because of the pilot. Decomposing this LATE for brother-brother pairs reveals that the younger brothers are extremely likely to conform and choose math-science (0.922) if their older brother is a complier and also quite likely to choose math-science (0.431) if the older brother does not choose math-science. This is consistent with a mechanism of sibling rivalry typically found to be systematically stronger among pairs of brothers. In addition, it may be interpreted as evidence against parental involvement driving the findings. One would think that parents would not systematically influence sons differently than daughters, unless they have a stereotypical mindset. We investigate this further in subsection 5.3.

Table 4. Estimates of Spillover Effects: Heterogeneity by Gender

	Parameter Estimates											
	Older Brother						**Older Sister**					
	Younger Brother			**Younger Sister**			**Younger Brother**			**Younger Sister**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
	First-stage: Older Sibling *MathScience*											
PilotIntro	0.128 ***	0.125 ***	0.118 ***	0.084 **	0.085 **	0.086 **	0.054 *	0.053 *	0.062 *	0.039	0.030	0.031
	(0.040)	(0.040)	(0.042)	(0.038)	(0.039)	(0.041)	(0.030)	(0.030)	(0.032)	(0.026)	(0.026)	(0.027)
Relative to overall first-stage	1.882	1.812	1.710	1.235	1.232	1.246	0.794	0.768	0.899	0.574	0.435	0.449
	Reduced-form: Younger Sibling *MathScience*											
PilotIntro	0.089 **	0.088 **	0.059	0.008	0.004	0.001	0.039	0.033	0.027	0.006	0.001	0.002
	(0.038)	(0.038)	(0.040)	(0.022)	(0.021)	(0.022)	(0.036)	(0.037)	(0.038)	(0.019)	(0.019)	(0.020)
	Outcome: Younger Sibling *MathScience*											
Older Sibling *MathScience*	0.696 **	0.703 **	0.504	0.093	0.053	0.017	0.732	0.629	0.430	0.164	0.037	0.051
	(0.306)	(0.315)	(0.329)	(0.254)	(0.247)	(0.257)	(0.738)	(0.738)	(0.620)	(0.489)	(0.633)	(0.624)
Control Variables:												
Entry Cohort Fixed Effects		+	+		+	+		+	+		+	+
County Indicators			+			+			+			+
Parental Variables (for Mother and Father):												
Highest Completed Education and Income		+	+		+	+		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+			+
Mean of Older Sibling *MathScience*		0.495			0.473			0.197			0.183	
Mean of Younger Sibling *MathScience*		0.313			0.089			0.307			0.093	
Number of Sibling Pairs		**1,710**			**1,963**			**1,780**			**2,333**	

Note: Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table 5. Estimates of Spillover Effects: Heterogeneity by Ability

	Parameter Estimates (Standard Errors)								
	Older Sibling								
	GPA < P50			P50 < GPA < P90			GPA > P90		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	First-stage: Older Sibling *MathScience*								
PilotIntro	0.086 ***	0.077 ***	0.071 ***	0.040	0.052 *	0.060 **	0.135 **	0.126 **	0.152 **
	(0.024)	(0.023)	(0.024)	(0.028)	(0.027)	(0.028)	(0.063)	(0.059)	(0.061)
Relative to overall first-stage	1.265	1.116	1.029	0.588	0.754	0.870	1.985	1.826	2.203
	Reduced-form: Younger Sibling *MathScience*								
PilotIntro	0.035 *	0.028	0.015	0.059 **	0.050 **	0.040 *	-0.051	-0.036	-0.043
	(0.019)	(0.019)	(0.019)	(0.024)	(0.023)	(0.024)	(0.051)	(0.050)	(0.052)
	Outcome: Younger Sibling *MathScience*								
Older Sibling *MathScience*	0.408	0.366	0.215	1.452	0.953	0.660	-0.376	-0.285	-0.285
	(0.250)	(0.257)	(0.282)	(1.066)	(0.600)	(0.462)	(0.419)	(0.390)	(0.325)
Control Variables:									
Entry Cohort Fixed Effects		+	+		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+		+	+
County Indicators			+			+			+
Parental Variables (for Mother and Father):									
Highest Completed Education and Income		+	+		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+
Mean of Older Sibling *MathScience*		0.290			0.347			0.420	
Mean of Younger Sibling *MathScience*		0.170			0.208			0.210	
Number of Sibling Pairs		**3,768**			**3,142**			**876**	

Note: Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

In Table 5 we split the sample by high school GPA of the older sibling. Unfortunately, we do not have any measures of ability (such as test scores or grades) before the students entered high school. Therefore, we define the subgroups based on GPA as measured as the simple average of grades in all high school courses at the end of high school. We caution that this measure may be affected by the course choice; for example, this GPA measure has more weight on performance in math-science related tasks for older siblings who chose more advanced math-science courses. However, this may be desirable for our particular purpose of measuring older sibling success in their chosen high school track. We distinguish three groups: below 50th percentile, between 50th-90th percentile, and above 90th percentile in the GPA distribution. The results show that the first stage is sizable as the unexpected pilot introduction increased the math-science probability by 13-15 percentage points for the older siblings with the highest ability, while the sibling spillover turns negative (but insignificant) in this case. This pattern of results supports the hypothesis of sibling competition. If the older sibling performs very well (is a "superstar") the younger sibling would rather not compete with this top performance and chooses another course combination. However, the younger sibling does conform to and compete with an older sibling with more moderate ability.[27]

[27] We have also run these regressions separately by gender, which reveals the same patterns, but the samples of high-ability students are only around 400 students in these cases and the estimates consequently very imprecise.

Table 6. Estimates of Spillover Effects: Heterogeneity by Parental Education

	Parameter (Standard Error)							
	Father				Mother			
	STEM		> 4-year college		STEM		> 4-year college	
	0	1	0	1	0	1	0	1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PilotIntro	0.074 ***	0.063 *	0.052 ***	0.135 ***	0.066 ***	0.106	0.062 ***	0.165 **
	(0.020)	(0.036)	(0.019)	(0.040)	(0.017)	(0.089)	(0.017)	(0.079)
Relative to overall first-stage	1.072	0.913	0.754	1.957	0.957	1.536	0.899	2.391
PilotIntro	0.016	0.038	0.010	0.063 *	0.013	0.166 **	0.011	0.168 ***
	(0.016)	(0.031)	(0.016)	(0.032)	(0.015)	(0.076)	(0.015)	(0.063)
Older Sibling *MathScience*	0.215	0.607	0.188	0.467 *	0.201	1.574	0.172	1.015 *
	(0.222)	(0.573)	(0.315)	(0.264)	(0.225)	(1.309)	(0.232)	(0.607)
Control Variables:								
Entry Cohort Fixed Effects	+	+	+	+	+	+	+	+
Sibling Pair Gender Composition	+	+	+	+	+	+	+	+
County Indicators	+	+	+	+	+	+	+	+
Parental variables (for mother and father):								
Highest Completed Education and Income	+	+	+	+	+	+	+	+
HS Mean og Highest Completed Education and Income	+	+	+	+	+	+	+	+
Mean of Older Sibling *MathScience*	0.302	0.406	0.321	0.349	0.323	0.398	0.326	0.354
Mean of Younger Sibling *MathScience*	0.172	0.242	0.183	0.210	0.186	0.245	0.188	0.214
Number of Sibling Pairs	**5,828**	**1,958**	**5,936**	**1,850**	**7,309**	**477**	**7,226**	**560**

Note: The STEM definition follows the definition by the National Science Foundation (NSF), which includes math, engineering, natural and technical sciences as well as some social sciences and life sciences. Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

5.2. Parental Background

Parents' and children's education are known to be strongly correlated for many reasons. One of the potential reasons is that parents' level and field of education reflect their ability to assist their children with making adequate education choices. Such assistance is highly relevant when it comes to choosing course packages in high school.

We do not know to what extent parents are actually involved in course choice in our sample. However, we do know that during high school when students are 16-20 years old, parents are still very closely involved in their educational decisions. Actually the family is ranked first when it comes to helping the students deciding what to do after compulsory school. A recent survey showed that 79% of students in 9-10th grade responded that their parents were closely involved in their high school enrollment choice. [28] We have no reason to believe that parental involvement is less in our sample.

This suggests that students discuss their education choices with their parents across socio-economic status. However, for parents without a college degree, it may be difficult to give appropriate advice about course choices. Parents with a STEM degree may be better at advising on courses in advanced math and science.

Stinebrickner and Stinebrickner (2014) model students' college major completion as the result of a learning process. They find that students tend to be particularly overoptimistic about whether they can complete a degree in math or science. The reason is that they have misperceptions about their ability to perform well in math or science. They suggest that such misperceptions are mainly prevalent among students who are less likely to have college-educated parents, because such students may be especially uninformed about what takes place during college. Arcidiacono, Aucejo, Fang and Spenner (2011) also find that students from ethnic groups who on average have less educated parents have more misperception about their graduation probabilities in natural sciences. We test whether a similar mechanism may play a role in advanced high school course choices.

In Table 6, we investigate heterogeneity of the incentive effects by parents' education. We distinguish students whose mother or father, respectively, has completed more than a 4-year college as well as students whose mother or the father has a STEM degree.

We find that the older sibling tends to respond more to the pilot scheme when either of the parents has completed a long college education or when the mother has a STEM degree. If the father has a STEM degree, the older sibling does not significantly react to the introductions of the pilot scheme.

[28] See Ministry of Education (2013).

This is most likely because the potential is exhausted as 40.6% of the high school students in such families already choose math and science (see the bottom panel of Table 6). When we divide by gender of the older sibling in Table B12, results suggest that older daughters respond more strongly to the pilot in families where at least one of the parents has a STEM degree, while the opposite is the case for older sons.[29] We interpret the complier analysis across parental education as suggestive evidence that information sharing in the family is important for course choices, but does not drive the sibling spillover. However, our estimates are not precise enough to draw sharp inferences.

The estimates of the spillover effects are rarely significant in the subsamples, but the point estimates are systematically higher when the parents have a longer college degree or a STEM degree. It is not evident that this necessarily reflects that peer effects are higher in these instances. It may reflect that the parents are involved in the course choice of both siblings. If the parents assist the older sibling with the course choice, they may transfer their acquired knowledge about the content and the demand of the courses to the younger sibling, and then the social interaction partly goes through the parents. This result also means that it is unlikely that the peer effect embodies the older siblings being direct path breakers by providing information on course content.

5.3. Long-term Outcomes

We have documented that there *is* a spillover effect of the pilot program on younger siblings' course choices; particularly for relatively closely spaced brothers. What remains is to examine whether there are any longer term effects on the younger siblings' completed education and earnings in their early careers.

Table 7 presents estimates of sibling peer effects on younger sibling's earnings at ages 30-35 and their highest completed education at age 30. Overall, there is some indication that younger siblings are 2-3 percentage points more likely to complete a long (> 4-year) college degree in a STEM field if the older sibling was unexpectedly exposed to the pilot scheme. Examining subsamples by younger sibling gender reveals that there are strong and significant positive spillovers on younger brothers, but no significant spillovers on younger sisters. This corroborates that there is not only a shorter-term effect of younger brothers being 5-6 percentage points more likely to choose advanced math-science if their older siblings unexpectedly got the option of choosing advanced math-science at a lower cost.

[29] This may be because there is a large pool of unexploited talent for older daughters in such families (only 25% choose math and science) and not for sons (54% of them choose math and science), who are likely to be always-takers. In Joensen and Nielsen (2016), we analyze the marginal monetary payoff for the students directly affected by the pilot. We find that the payoff tends to be high for those who gain the most, while it approaches zero for boys when more than half of them choose math and science.

Table 7. Estimates of Spillover Effects: Long-term Outcomes

	Parameter Estimates (Standard Errors)								
	All			Younger Sibling					
				Brother			Sister		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	First-stage: Older Sibling *MathScience*								
PilotIntro	0.067 ***	0.068 ***	0.068 ***	0.072 ***	0.076 ***	0.078 ***	0.062 **	0.058 **	0.057 **
Younger Sibling, age 30	(0.018)	(0.017)	(0.018)	(0.027)	(0.026)	(0.027)	(0.024)	(0.023)	(0.024)
Relative to overall first-stage				1.059	1.101	1.130	0.912	0.841	0.826
	Reduced-form: Younger Sibling longer-term outcomes								
PilotIntro	0.088	0.068	0.035	0.254 **	0.235 *	0.242 *	-0.062	-0.057	-0.103
Younger Sibling log(earnings), ages 30-35	(0.089)	(0.088)	(0.093)	(0.121)	(0.122)	(0.130)	(0.126)	(0.126)	(0.132)
PilotIntro	0.007	0.012	0.011	0.054 *	0.063 **	0.068 **	-0.034	-0.030	-0.035
Younger Sibling > 4-year college, age 30	(0.019)	(0.018)	(0.019)	(0.028)	(0.028)	(0.029)	(0.025)	(0.024)	(0.025)
PilotIntro	0.025	0.024	0.024	0.063 **	0.068 **	0.066 **	-0.012	-0.008	-0.005
Younger Sibling STEM field college, age 30	(0.017)	(0.017)	(0.017)	(0.027)	(0.027)	(0.028)	(0.021)	(0.021)	(0.021)
PilotIntro	0.023	0.025	0.026 *	0.056 **	0.063 ***	0.064 **	-0.008	-0.004	-0.002
Younger Sibling > 4-year college in STEM field, age 30	(0.015)	(0.015)	(0.016)	(0.025)	(0.024)	(0.025)	(0.019)	(0.019)	(0.019)
	Outcome: Younger Sibling longer-term outcomes								
Older Sib *MathScience*	1.232	0.942	0.492	3.129 *	2.811 *	2.872	-1.009	-0.971	-1.756
Younger Sibling log(earnings), ages 30-35	(1.272)	(1.231)	(1.285)	(1.824)	(1.665)	(1.756)	(2.120)	(2.212)	(2.423)
Older Sib *MathScience*	0.106	0.174	0.163	0.749	0.825 *	0.883 *	-0.554	-0.505	-0.621
Younger Sibling > 4-year college, age 30	(0.276)	(0.267)	(0.274)	(0.463)	(0.443)	(0.455)	(0.462)	(0.464)	(0.514)
Older Sib *MathScience*	0.371	0.359	0.356	0.878 *	0.896 **	0.853 *	-0.190	-0.137	-0.091
Younger Sibling STEM field college, age 30	(0.266)	(0.254)	(0.261)	(0.484)	(0.448)	(0.442)	(0.350)	(0.362)	(0.377)
Older Sib *MathScience*	0.340	0.364	0.389	0.784 *	0.832 **	0.825 **	-0.130	-0.071	-0.041
Younger Sibling > 4-year college in STEM field, age 30	(0.240)	(0.233)	(0.242)	(0.437)	(0.411)	(0.411)	(0.315)	(0.324)	(0.340)
Control Variables:									
Entry Cohort Fixed Effects		+	+		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+		+	+
County Indicators			+			+			+
Parental Variables (for Mother and Father):									
Highest Completed Education and Income		+	+		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+
Mean of Outcome variables:									
Younger Sib. log(earnings), ages 30-35		11.076			11.275			10.918	
Younger Sib. > 4-year college, age 30		0.357			0.399			0.324	
Younger Sib. STEM field college, age 30		0.268			0.359			0.196	
Younger Sib. > 4-year college in STEM field, age 30		0.197			0.248			0.156	
Number of Sibling Pairs, Younger Sib. obs. at age 30		**7,123**			**3,151**			**3,972**	
Number of Observations, Younger Sib. earnings ages 30-35		**42,484**			**18,787**			**23,697**	

Note: The definition of STEM fields follows the definition by the National Science Foundation (NSF), which includes math, engineering, natural and technical sciences as well as some social sciences and life sciences. . Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

The reduced-form estimates in Table 7 show that these younger brothers also become 5-7 percentage points more likely to complete a long college degree, 6-7 percentage points more likely to complete a (long) STEM-field college degree, and that they earn 24-25 log-points more when they are age 30-35 years old.

From previous studies we know that there is a substantial direct effect of investing in math on career outcomes of the individual (Cortes et al. 2015; Falch et al. 2014; Joensen and Nielsen, 2009; 2016). Based on our analysis of siblings, we now also know that this investment spills over to younger siblings and that this indeed has long-term consequences for career outcomes of the younger siblings.

6. Conclusion

We study the spillover effects of an education program offered to older siblings on their younger siblings' education choices. We find significant spillover effects, which are particularly strong for brothers, and present evidence that social interactions among siblings are important for skill formation.

Younger siblings are 2-3 percentage points more likely to choose advanced math and science in high school if their older siblings could unexpectedly opt for this course choice at a reduced cost. If we invoke the exclusion restriction that the extra education program offered to the older sibling only influences the younger sibling through the older sibling's course choice, this implies a sizeable peer effect as almost half of the direct effect spills over, which varies across subgroups. Spillovers are never significant for sisters and strongest for relatively closely spaced brother-brother pairs. We argue that the pattern of the heterogeneous spillover effects across gender, ability, birth spacing and birth order suggests that the most likely mechanism is competition and sibling rivalry. Furthermore, we document that these sibling spillovers have long-term consequences for the careers of the affected subgroup of younger brothers.

Our results indicate that social interactions in the family may exacerbate inequality across households. First, the presence of possibly large sibling peer effects in education choices reveals that the strong sibling correlations could conceal influential interactions among siblings. Second, we find systematically larger spillover effects in families where either of the parents has a long college education or a STEM degree. This suggests that parents are closely involved in the educational decision and that they are part of the social environment in which these decisions are made. However, our empirical analysis of the potential mechanisms are merely suggestive and a first step towards sheding light on what the sibling spillovers in education choices embody. More hard evidence is needed on the role of the family and siblings in influencing eachother's educational decisions and human capital investments more generally.

References

Abadie, A. (2002), Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models. *Journal of the American Statistical Association* 97(457): 284-292.

Abadie, A. (2003), Semiparametric Instrumental Variable Estimation of Treatment Response Models. *Journal of Econometrics* 113: 231 – 263.

Adams, B. N. (1972), Birth Order: A Critical Review. *Sociometry* 35: 411-439.

Adler, A. (1927), *Understanding human nature*. New York: Greenburg.

Akerlof, G. (1997), Social Distance and Social Decisions. *Econometrica* 65(5): 1005-1027.

Akerlof, G. and R. Kranton (2002), Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature* 40(4): 1167-1201.

Altonji, J. (1995), The Effect of High School Curriculum on Education and Labor Market Outcomes. *Journal of Human Resources* 30(3): 409-438.

Altonji, J. G., E. Blom and C. Meghir (2012), Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. *Annual Review of Economics* 4: 185-223.

Altonji, J. G., S. Cattan and I. Ware (2017), Identifying Sibling Influence on Teenage Substance Use. *Journal of Human Resources* 52(1): 1-47.

Altonji, J. G., T. Elder and C. Taber (2005), An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling. *Journal of Human Resources* 40(4):791-821.

Arcidiacono, P., E. M. Aucejo, H. Fang and K. I. Spenner (2011), Does affirmative action lead to mismatch? A new test and evidence. *Quantitative Economics* 2(3): 303-333.

Avvisati, F., M. Gurgand, N. Guyon and E. Maurin (2014), Getting Parents Involved: A Field Experiment in Deprived Schools. *Review of Economic Studies* 81(1): 57-83.

Becker, G. S. and N. Tomes (1979), An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87(6): 1153-1189.

Becker, G. S. and N. Tomes (1986), Human capital and the rise and fall of families. *Journal of Labor Economics* 4(3): S1-S39.

Bingley, P., P. Lundborg and S. V. Lyk-Jensen (2017), Brothers in Arms: Spillovers from a Draft Lottery, IZA Discussion Paper No. 10483.

Björklund, A. and M. Jäntti (2012), How important is family background for labor-economic outcomes? *Labour Economics*, 19(4): 465-474.

Björklund, A. and K. G. Salvanes (2010), Education and family background: mechanisms and policies. In: Hanushek, E.A., S. Machin and L. Woessman (Eds.), *Handbooks in Economics of Education*, Vol. 3. North-Holland: Amsterdam, 201-247.

Black, S. E., S. Breining, D. Figlio, J. Guryan, K. Karbownik, H. S. Nielsen, J. Roth and M. Simonsen (2017), Sibling Spillovers. NBER Working Paper No. 23062.

Black, S. E. and P. J. Devereux (2010), Recent developments in intergenerational mobility. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, North-Holland: Amsterdam, 1487-1541.

Black, S. E., P. J. Devereux and K. G. Salvanes (2005), The more the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics* 120: 669-700.

Breining, S., N. M. Daysal, M. Simonsen and M. Trandafir (2015), Spillover Effects of Early-Life Medical Interventions. IZA Discussion Paper No. 9086.

Buhrmester, D. (1992), The Developmental Courses of Sibling and Peer Relationships, in F. Boer and J. Dunn (Eds.), *Children's Sibling Relationships: Developmental and Clinical Issues*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Buser, T., M. Niederle and H. Oosterbeek (2014), Gender, competitiveness and career choices. *Quarterly Journal of Economics* 129(3): 1409-1447.

Card, D. and L. Giuliano (2013), Peer Effects and Multiple Equilibria in the Risky Behavior of Friends. *Review of Economics and Statistics* 95(4): 1130-1149.

Conley, D. (2000), Sibship, Sex Composition and the Educational Attainment of Men and Women. *Social Science Research* 29: 441-457.

Cortes, K., J. Goodman and T. Nomi (2015), Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. *Journal of Human Resources* 50(1): 108-158.

Cunha, F. and J. J. Heckman (2007), The technology of skill formation. *American Economic Review* 97(2): 31-47.

Cunha, F., Heckman, J. J. and S. Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, *57*(2): 191-261.

Cunha, F., J. J. Heckman and S. M. Schennach (2010), Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78 (3): 883-931.

Dahl, G., A. R. Kostøl and M. Mogstad (2014), Family Welfare Cultures. *Quarterly Journal of Economics* 129 (4): 1711-1752.

Dahl, G., K. V. Løken and M. Mogstad (2014), Peer Effects in Program Participation. *American Economic Review* 104(7): 2049-74.

Falch, T., O. H. Nyhuus and B. Strøm (2014), Causal Effect of Mathematics. *Labour Economics.* 31: 174-187.

Glaeser, E. L., B. I. Sacerdote and J. A. Scheinkman (2003), The Social Multiplier, *Journal of the European Economic Association* 1(2-3): 345-353.

Grose, M. (1991), Why First-Borns Rule the World and Last-Borns Want to Change it. Random House: Australia.

Heckman, J. J. and S. Mosso (2014), The Economics of Human Development and Social Mobility. *Annual Review of Economics* 6: 689-733.

Huggett, M., G. Ventura and A. Yaron (2011). Sources of Lifetime Inequality. *American Economic Review*, 101(7): 2923-54.

Joensen, J. S. and H. S. Nielsen (2009), Is there a Causal Effect of High School Math on Labor Market Outcomes? *Journal of Human Resources* 44(1): 171-198.

Joensen, J. S. and H. S. Nielsen (2016), Mathematics and Gender: Heterogeneity in Causes and Consequences. *Economic Journal* 126: 1129-1163.

Keane, M. P. and K. I. Wolpin (1997), The Career Decisions of Young Men. *Journal of Political Economy*, 105(3): 473-522.

Manski, C. F. (1993), Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60(3): 531-542.

Manski, C. (1995), *Estimation Problems in the Social Sciences*. Cambridge, Massachusetts: Harvard University Press.

Mazumder, B. (2008), Sibling similarities and economic inequality in the US. *Journal of Population Economics* 21: 685-701.

Ministry of Education (2013), UU-centrenes vejledning (In English: Advice from the Centers for Youth Education). UNI-C, Ministry of Education.

Moffitt, R. A. (2001), Policy interventions, low-level equilibria, and social interactions. In S. N. Durlauf and H. P. Young (eds.), *Social Dynamics*. Cambridge: MIT Press, 45-82.

Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122(3): 1067-1101.

Niederle, M. and L. Vesterlund (2010). Explaining the gender gap in Math test scores: the role of competition, *Journal of Economic Perspectives*, 24(2): 129-144.

Qureshi, J. (2016), Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital. Manuscript, University of Illinois at Chicago.

Stinebrickner, R. and T. Stinebrickner (2014), A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *Review of Economic Studies* 81(1): 426-472.

Appendix A. Earnings across Course Choices

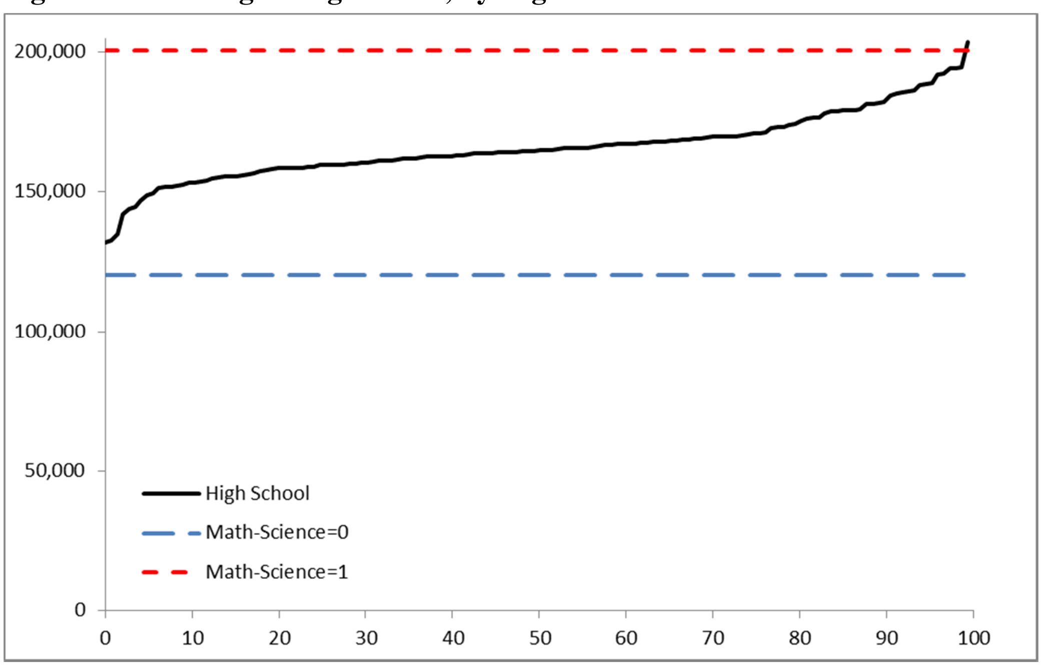
In the empirical analysis, we focus on the choice of advanced math and science. This course combination gives access to STEM field colleges programs without further supplementary courses, and on average, the students choosing this course combination have the most lucrative careers.

Figure A1 shows that the choice of advanced math-science increases earnings by more than moving from the "worst" to the "best" high school, where "worst" ("best") refers to lowest (highest) ranked high school as measured by the average earnings of graduates. Average earnings for non-math-science high school graduates are lower than the earnings at the "worst" school, while average earnings of math-science graduates are as high as top ranked schools.

Figures A2 and A3 show that the average earnings premium of advanced math-science is higher than the average earnings premium of non-math-science high school graduates relative to those without a traditional academic high school degree (black solid line). Choosing advanced math-science thus seems to increase earnings by more than acquiring a traditional academic high school degree per se. This is true both at age 30 (Figure A2) and age 35 (Figure A3).

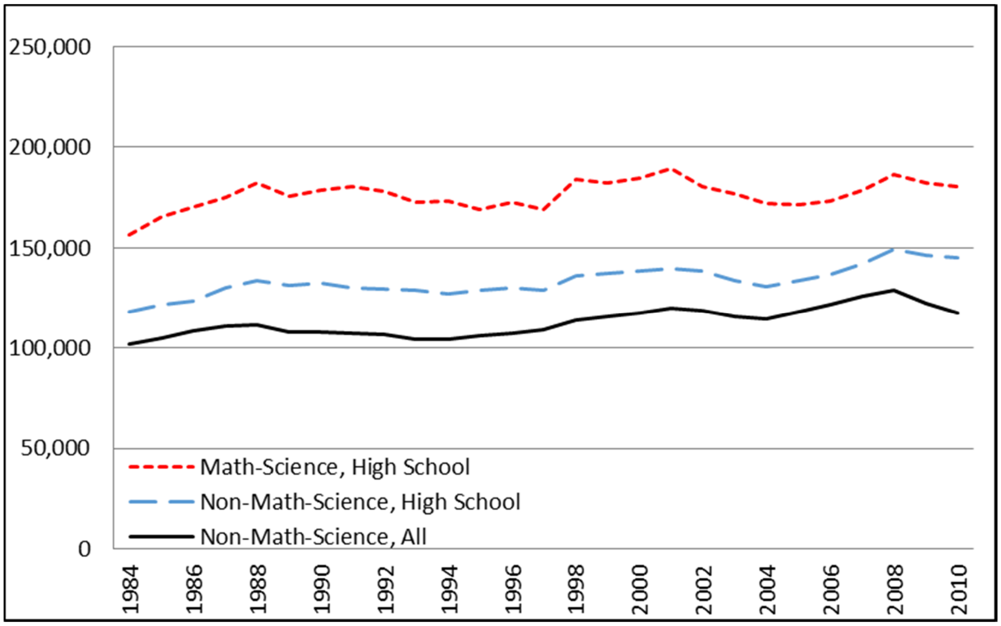
Thus Figures A1-A3 present descriptive evidence that the choice of advanced math-science is more strongly related to earnings at ages 30-35 than the choice of high school.

Figure A1. Earnings at Ages 30-35, by High School and Advanced Course Choices



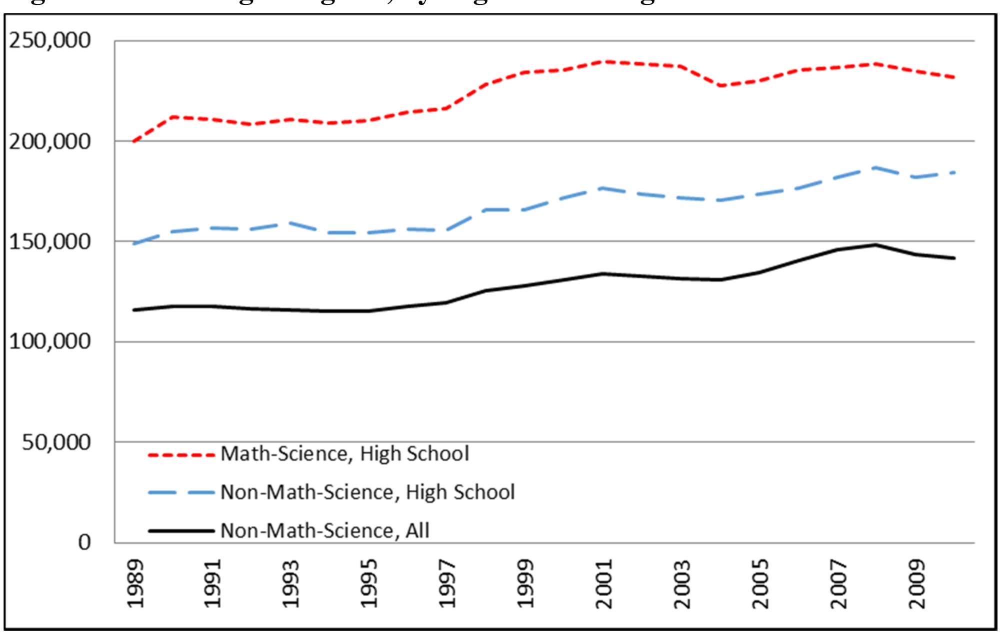
Note: The figure displays annual earnings of high school graduates for the each cohort at age 30-35. Annual earnings are displayed separately by high school and advanced Math-Science. The horizontal axis displays the percentile rank of the high schools in our sample based on average earnings for all students graduating from the high school (black solid line). All numbers are in real 1989 DKK. The exchange rate was 6.1853 USD/DKK ultimo December 1989.

Figure A2. Earnings at age 30, by High School Degree and Advanced Course Choices



Note: The figure displays annual earnings of each cohort who is 30 years old in the relevant year: 1984-2010. Thus most of those observed in 2001 enrolled in high school in 1987. Annual earnings are displayed separately by high school degree and advanced Math-Science choice. All numbers are in real 1989 DKK. The exchange rate was 6.1853 USD/DKK ultimo December 1989.

Figure A3. Earnings at age 35, by High School Degree and Advanced Course Choices



Note: The figure displays annual earnings of each cohort who is 35 years old in the relevant year: 1989-2010. Thus most of those observed in 2006 enrolled in high school in 1987. Annual earnings are displayed separately by high school degree and advanced Math-Science. All numbers are in real 1989 DKK. The exchange rate was 6.1853 USD/DKK ultimo December 1989.

Appendix B. Additional Descriptive Statistics and Results

Table B1. Overview of Sample Selection

	N	MathScience
All high school entry cohorts 1984-87	79,681	0.2192
who graduates within 3 years at ages 18-22	64,489	0.2594
with younger siblings		
sibling pairs	55,734	0.2655
accounting for older siblings once	38,273	0.2706
with younger siblings in high school		
sibling pairs	28,509	0.3008
accounting for older siblings once	23,046	0.2991
with younger siblings in high school*		
sibling pairs	24,116	0.3075
accounting for older siblings once	20,016	0.3055
with younger siblings in high school cohorts 1988-97*		
sibling pairs	18,780	0.3088
accounting for older siblings once	16,313	0.3074
and age difference < 10 years*		
sibling pairs	**17,691**	0.3115
accounting for older siblings once	15,420	0.3102
with younger siblings in high school cohorts 1988-91*		
sibling pairs	12,157	0.3201
accounting for older siblings once	11,610	0.3183
and age difference < 4 years*		
sibling pairs	**7,786**	0.3279
accounting for older siblings once	7,496	0.3280

* Graduated with in 3 years at ages 18-22.

Table B2. Summary of Older Siblings' Course Choice,
by their Exposure to the Pilot Scheme and by High School Cohort of Younger Sibling

Younger Sib High School Cohort	**Older Sibling** Pilot School = 0 Pilot Intro = 0		**Older Sibling** Pilot School = 0 Pilot Intro = 1		**Older Sibling** Pilot School = 1 Pilot Intro = 0		All	
	N	$MathScience_{old}$	N	$MathScience_{old}$	N	$MathScience_{old}$	N	$MathScience_{old}$
1988	2,160	0.2815	443	0.3341	689	0.3614	3,292	0.3053
1989	2,192	0.2810	431	0.3364	953	0.4292	3,576	0.3272
1990	1,788	0.2768	331	0.4018	819	0.4371	2,938	0.3356
1991	1,384	0.2724	257	0.2763	710	0.3972	2,351	0.3105
1992	1,023	0.2366	214	0.3131	500	0.3740	1,737	0.2855
1993	858	0.2669	155	0.3677	392	0.3648	1,405	0.3053
1994	669	0.2272	123	0.2764	326	0.3926	1,118	0.2809
1995	614	0.2492	115	0.3391	260	0.4154	989	0.3033
1996	452	0.2235	93	0.2796	222	0.3739	767	0.2738
1997	391	0.2532	73	0.3014	143	0.2727	607	0.2636
Total	11,531	0.2664	2,235	0.3320	5,014	0.3961	18,780	0.3088
1988-1997	10,733	0.2679	2,097	0.3348	4,861	0.3979	17,691	0.3115
1988-1991	4,463	0.2772	865	0.3457	2,458	0.4138	7,786	0.3279

Note: The table displays the number of younger siblings and the fraction of their older siblings choosing advanced Math with advanced Physics or Chemistry. The numbers are displayed by younger siblings' high school entry cohort and type of high school attended by the older sibling: traditional high schools only offering advanced Math with advanced Physics (*PilotSchool*=0, *PilotIntro*=0), schools which unexpectedly introduced the pilot scheme combining advanced Math with advanced Chemistry (*PilotSchool*=0, *PilotIntro*=1) for the older sibling, and schools which already had adopted the pilot scheme (*PilotSchool*=1, *PilotIntro*=0). The two rows at the bottom summarize information for younger sibling enrolling high school 1988-97 (age gap ≤ 10 years) and 1988-91 (age gap ≤ 4 years).

Table B3. Younger Sibling Course Choice, by Course Choice of Older Sibling

Younger Sibling		Older Sibling MathScience = 0		Older Sibling MathScience = 1		All	
Gender	Course Choice	N	Mean	N	Mean	N	Mean
		All					
All	MathScience	5,233	0.1439	2,553	0.2832	7,786	0.1896
		Older Brother					
Brother	MathScience	863	0.2063	847	0.4227	1,710	0.3135
Sister	MathScience	1,035	0.0444	928	0.1390	1,963	0.0891
		Older Sister					
Brother	MathScience	1,430	0.2804	350	0.4171	1,780	0.3073
Sister	MathScience	1,905	0.0672	428	0.2103	2,333	0.0934

Note: The table displays math-science course choices of younger siblings, by sibling pair gender composition and by older sibling's math-science choice.

Table B4. Younger Sibling Course Choice, by Pilot School Status of Older Sibling

Younger Sibling		Older Sibling Pilot School = 0 Pilot Intro = 0		Older Sibling Pilot School = 0 Pilot Intro = 1		Older Sibling Pilot School = 1 Pilot Intro = 0		All	
Gender	Course Choice	N	Mean	N	Mean	N	Mean	N	Mean
		All							
All	MathScience	4,463	0.1804	865	0.2162	2,458	0.1969	7,786	0.1896
		Older Brother							
Brother	MathScience	1,001	0.3027	189	0.3915	520	0.3058	1,710	0.3135
Sister	MathScience	1,094	0.0804	204	0.0882	665	0.1038	1,963	0.0891
		Older Sister							
Brother	MathScience	1,005	0.2955	212	0.3349	563	0.3179	1,780	0.3073
Sister	MathScience	1,363	0.0858	260	0.0923	710	0.1085	2,333	0.0934

Note: The table displays math-science course choices of younger siblings. The numbers are displayed by sibling pair gender composition and type of high school attended by the older sibling: traditional high schools only offering advanced Math with advanced Physics (*PilotSchool*=0, *PilotIntro*=0), schools which unexpectedly introduced the pilot scheme combining advanced Math with advanced Chemistry (*PilotSchool*=0, *PilotIntro*=1) for the older sibling, and schools which already had adopted the pilot scheme (*PilotSchool*=1, *PilotIntro*=0).

Table B5. Descriptive Statistics, by Pilot School Status of Older Sibling

	Older Sibling Pilot School = 0 Pilot Intro = 0		**Older Sibling** Pilot School = 0 Pilot Intro = 1			**Older Sibling** Pilot School = 1 Pilot Intro = 0		
Younger Sibling	Mean	Std.Dev.	Mean	Std.Dev.	MeanDiff	Mean	Std.Dev.	MeanDiff
Male	0.4495	0.4975	0.4636	0.4990	-0.0141	0.4406	0.4966	0.0089
High school cohort								
1988	0.4098	0.4919	0.4324	0.4957	-0.0226	0.2738	0.4460	**0.1360**
1989	0.3274	0.4693	0.3110	0.4632	0.0164	0.3515	0.4775	**-0.0241**
1990	0.1927	0.3945	0.1746	0.3798	0.0181	0.2433	0.4292	**-0.0506**
1991	0.0701	0.2554	0.0821	0.2746	-0.0119	0.1314	0.3379	**-0.0613**
Father								
Log(Earnings)	10.1093	4.5365	10.1125	4.5967	-0.0032	10.2757	4.4347	-0.1664
Primary School Only	0.1214	0.3267	0.1422	0.3495	*-0.0208*	0.1098	0.3128	0.0116
High School Only	0.0432	0.2034	0.0497	0.2175	-0.0065	0.0439	0.2050	-0.0007
Vocational Training	0.2543	0.4355	0.2509	0.4338	0.0034	0.2327	0.4226	**0.0216**
2-year College	0.0356	0.1854	0.0451	0.2076	-0.0095	0.0374	0.1898	-0.0018
4-year College	0.2642	0.4409	0.2474	0.4318	0.0168	0.2628	0.4403	0.0014
> 4-year College	0.2252	0.4178	0.2116	0.4087	0.0136	0.2604	0.4389	**-0.0352**
STEM Field	0.2389	0.4264	0.2532	0.4351	-0.0143	0.2738	0.4460	**-0.0349**
Mother								
Log(Earnings)	9.5384	4.3456	9.5306	4.3265	0.0078	9.6555	4.2569	-0.1171
Primary School Only	0.1499	0.3570	0.1526	0.3598	-0.0027	0.1318	0.3384	**0.0181**
High School Only	0.0379	0.1909	0.0347	0.1831	0.0032	0.0391	0.1938	-0.0012
Vocational Training	0.3034	0.4598	0.3179	0.4659	-0.0145	0.3096	0.4624	-0.0062
2-year College	0.0614	0.2401	0.0751	0.2638	-0.0138	0.0736	0.2612	**-0.0122**
4-year College	0.3424	0.4746	0.3306	0.4707	0.0117	0.3393	0.4736	0.0031
> 4-year College	0.0726	0.2595	0.0578	0.2335	0.0148	0.0732	0.2606	-0.0006
STEM Field	0.0596	0.2368	0.0497	0.2175	0.0099	0.0683	0.2524	-0.0087
Older sibling								
Male	0.4694	0.4991	0.4543	0.4982	0.0151	0.4821	0.4998	-0.0127
Math Science	0.2772	0.4477	0.3457	0.4759	**-0.0685**	0.4138	0.4926	**-0.1366**
GPA	8.5817	0.9721	8.5032	0.9260	**0.0784**	8.5680	0.9679	0.0137
First Born	0.7930	0.4052	0.8000	0.4002	-0.0070	0.7815	0.4133	0.0114
Second Born	0.1667	0.3728	0.1607	0.3675	0.0060	0.1798	0.3841	-0.0131
Third Born or Later	0.0439	0.2049	0.0405	0.1972	0.0035	0.0419	0.2004	0.0020
Sibship								
Sibship Size = 2	0.4779	0.4996	0.5017	0.5003	-0.0238	0.4870	0.4999	-0.0091
Sibship Size = 3	0.3477	0.4763	0.3561	0.4791	-0.0083	0.3657	0.4817	-0.0180
Sibship Size = 4+	0.1743	0.3794	0.1422	0.3495	**0.0321**	0.1473	0.3545	**0.0270**
Number of Individuals	**4,463**		**865**			**2,458**		

Note: The table displays means and standard deviations of background variables for older siblings and their families by type of high school attended by the older sibling: traditional high schools only offering advanced Math with advanced Physics (*PilotSchool*=0, *PilotIntro*=0), schools which unexpectedly introduced the pilot scheme combining advanced Math with advanced Chemistry (*PilotSchool*=0, *PilotIntro*=1), and schools which had adopted the pilot scheme when the older sibling enrolled (*PilotSchool*=1, *PilotIntro*=0). The population includes 7,786 sibling pairs with an age gap ≤ 4 years where the younger sibling enrolled in high school in 1988-91. Bold and italics indicate that the mean is significantly different from the mean for Pilot School=0 & Pilot Intro=0 at the 5 % and the 10% level, respectively.

Table B6. Estimates of Spillover Effects: Probit Estimators

	Parameter Estimates (Standard Errors) [Average Marginal Effects]					
	Probit Normal Index Models					
	Probit			**Bivariate Probit**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	First-stage: Older Sibling *MathScience*					
PilotIntro				0.203 ***	0.222 ***	0.223 ***
				(0.046)	(0.050)	(0.052)
				[0.072]	[0.071]	[0.071]
	Outcome: Younger Sibling *MathScience*					
Older Sibling *MathScience*	0.493 ***	0.565 ***	0.566 ***	1.453 ***	0.793 **	0.655 *
	(0.034)	(0.038)	(0.039)	(0.350)	(0.321)	(0.355)
	[0.140]	[0.150]	[0.149]	[0.360]	[0.188]	[0.154]
Control Variables:						
Entry Cohort Fixed Effects		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+
County Indicators			+			+
Parental Variables (for Mother and Father):						
Highest Completed Education and Income		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+
Mean of Older Sibling *MathScience*	0.328					
Mean of Younger Sibling *MathScience*	0.190					
Number of Sibling Pairs	**7,786**					

Note: Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table B7. Means by Treatment Status

	Means			
	All	**Never Takers**	**Always Takers**	**Compliers**
Father				
log(earnings)	12.070	12.130	12.090	11.492
Primary School Only	0.132	0.149	0.118 **	0.021
High School Only	0.047	0.054	0.044	-0.013
Vocational Training	0.268	0.296 *	0.243 **	0.095
2-year College	0.039	0.041	0.042	0.012
4-year College	0.276	0.251	0.310 ***	0.369
> 4-year College	0.239	0.209 *	0.244	0.515
STEM Field	0.241	0.237	0.307 ***	0.016
Mother				
log(earnings)	11.440	11.380	11.460	12.017
Primary School Only	0.155	0.168	0.157	0.026
High School Only	0.039	0.035	0.047	0.040
Vocational Training	0.316	0.336	0.318	0.112
2-year College	0.066	0.082	0.057	-0.058
4-year College	0.351	0.332	0.347	0.553
> 4-year College	0.074	0.047 ***	0.073	0.326 ***
STEM Field	0.058	0.042 *	0.068	0.167
Siblings				
Older Sibling Male	0.467	0.325	0.723 ***	0.787 ***
Younger Sibling Male	0.452	0.445	0.472	0.432
Older Sibling's GPA	8.569	8.472 **	8.714 ***	8.909
Number of Observations (to Compute Means)	**5,328**	**566**	**1,237**	**-**
Number of Observations	5,328	3,486	1,477	364

Note: Table reports means of family background for individuals with *PilotSchool*=0. "Never Takers" are individuals who do not choose math-science even when exposed to the pilot program, i.e. *PilotIntro*=1 and *MathScience*=0; "Always Takers" are individuals who choose math-science even without the pilot program, i.e. *PilotIntro*=0 and *MathScience*=1; "Compliers" are individuals who only choose math-science when exposed to the pilot program, but not otherwise. The means for compliers are inferred based on the other columns. Standard errors are bootstrapped (99 repetitions) and significant mean differences between "All" high school students at non-pilot schools and those who unexpectedly introduce the pilot; i.e. *PilotIntro*=1, at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table B8. Estimates of Spillover Effects: Exclude Cohorts Close to 1988-Reform.

	Parameter Estimates (Standard Errors)					
	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	First-stage: Older Sibling *MathScience*					
PilotIntro				0.086 ***	0.080 ***	0.067 ***
Younger Sibling 1988-91, ≤4y, Older Sibling 1984-86				(0.021)	(0.020)	(0.021)
PilotIntro				0.068 ***	0.067 ***	0.069 ***
Younger Sibling 1989-91, ≤4y, Older Sibling 1984-87				(0.023)	(0.022)	(0.023)
	Reduced-form: Younger Sibling *MathScience*					
PilotIntro	0.040 **	0.031 *	0.029			
Younger Sibling 1988-91, ≤4y, Older Sibling 1984-86	(0.018)	(0.017)	(0.018)			
PilotIntro	0.049 **	0.052 ***	0.045 **			
Younger Sibling 1989-91, ≤4y, Older Sibling 1984-87	(0.020)	(0.018)	(0.019)			
	Outcome: Younger Sibling *MathScience*					
Older Sibling *MathScience*	0.144 ***	0.146 ***	0.145 ***	0.459 **	0.383 *	0.437
Younger Sibling 1988-91, ≤4y, Older Sibling 1984-86	(0.013)	(0.014)	(0.014)	(0.217)	(0.217)	(0.277)
Older Sibling *MathScience*	0.141 ***	0.141 ***	0.140 ***	0.728 **	0.785 **	0.654 **
Younger Sibling 1989-91, ≤4y, Older Sibling 1984-87	(0.012)	(0.012)	(0.012)	(0.348)	(0.342)	(0.321)
Control Variables:						
Entry Cohort Fixed Effects		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+
County Indicators			+			+
Parental Variables (for Mother and Father):						
Highest Completed Education and Income		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+

Note: The samples consist of 4,705 and 4,910 sibling pairs, respectively. Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table B9. Estimates of Spillover Effects: Exclude Closely Spaced Siblings

	Parameter Estimates (Standard Errors)					
	OLS			**2SLS**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	First-stage: Older Sibling *MathScience*					
PilotIntro				0.077 ***	0.077 ***	0.075 ***
Younger Sibling 1988-91, ≤4y				(0.019)	(0.018)	(0.018)
	Reduced-form: Younger Sibling *MathScience*					
PilotIntro	0.039 **	0.034 **	0.023			
Younger Sibling 1988-91, ≤4y	(0.016)	(0.015)	(0.016)			
	Outcome: Younger Sibling *MathScience*					
Older Sibling *MathScience*	0.146 ***	0.150 ***	0.150 ***	0.511 **	0.435 **	0.314
	(0.011)	(0.011)	(0.011)	(0.223)	(0.203)	(0.211)
Control Variables:						
Entry Cohort Fixed Effects		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+
County Indicators			+			+
Parental Variables (for Mother and Father):						
Highest Completed Education and Income		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+
Number of Sibling Pairs	**6,515**					

Note: Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table B10. Estimates of Spillover Effects: Distinguish Low versus High Fraction Math-Science Prior to Pilot

	Parameter Estimates (Standard Errors)											
	High School MathScience						**County MathScience**					
	< Median			**> Median**			**< Median**			**> Median**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
	First-stage: Older Sibling *MathScience*											
PilotIntro	0.082 ***	0.077 ***	0.062 **	0.069 ***	0.074 ***	0.051 *	0.044 *	0.051 **	0.036	0.099 ***	0.088 ***	0.102 ***
	(0.024)	(0.023)	(0.026)	(0.026)	(0.025)	(0.027)	(0.024)	(0.022)	(0.023)	(0.027)	(0.025)	(0.027)
Relative to Overall First-Stage	1.206	1.116	0.899	1.015	1.072	0.739	0.647	0.739	0.522	1.456	1.275	1.478
	Reduced-form: Younger Sibling *MathScience*											
PilotIntro	0.015	0.009	-0.020	0.058 **	0.058 ***	0.036	0.029	0.029	0.023	0.040 *	0.027	0.024
	(0.020)	(0.020)	(0.022)	(0.023)	(0.022)	(0.023)	(0.021)	(0.020)	(0.020)	(0.022)	(0.022)	(0.023)
	Outcome: Younger Sibling *MathScience*											
Older Sibling *MathScience*	0.183	0.117	-0.318	0.840 **	0.781 **	0.703	0.647	0.583	0.643	0.402 *	0.308	0.234
	(0.243)	(0.252)	(0.403)	(0.425)	(0.359)	(0.528)	(0.535)	(0.428)	(0.648)	(0.229)	(0.241)	(0.220)
Control Variables:												
Entry Cohort Fixed Effects		+	+		+	+		+	+		+	+
County Indicators			+			+			+			+
Parental Variables (for Mother and Father):												
Highest Completed Education and Income		+	+		+	+		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+			+
Mean of Older Sibling *MathScience*		0.317			0.340			0.326			0.330	
Mean of Younger Sibling *MathScience*		0.194			0.187			0.201			0.178	
Number of Sibling Pairs		**3,901**			**3,846**			**4,027**			**3,759**	

Note: The left (right) hand side of table distinguishes high schools (counties) where the fraction choosing math and science in 1983 was below and above the median. Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively. Note that we do not observe the course choices at one high school in 1983. This school was founded in 1985 and 39 older siblings in our sample subsequently attend this high school that never adopted the pilot program.

Table B11. Estimates of Spillover Effects: Randomly Matched Sibling Pairs

	Parameter Estimates (Standard Errors)											
	Only Children Matched by Mother's Income, Education level, and STEM field											
	Random High School						Same High School					
	OLS			2SLS			OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	First-stage: Older "Sibling" *MathScience*											
PilotIntro				0.105 ***	0.107 ***	0.091 **				0.087 **	0.092 **	0.067
				(0.036)	(0.034)	(0.036)				(0.042)	(0.041)	(0.043)
	Reduced-form: Younger "Sibling" *MathScience*											
PilotIntro	0.002	0.003	0.000				-0.004	0.010	-0.004			
	(0.031)	(0.031)	(0.032)				(0.038)	(0.037)	(0.039)			
	Outcome: Younger "Sibling" *MathScience*											
Older "Sibling" *MathScience*	0.017	0.020	0.023	0.014	0.032	-0.003	0.033	0.044	0.041	-0.042	0.113	-0.063
	(0.022)	(0.022)	(0.022)	(0.299)	(0.283)	(0.348)	-0.027	(0.027)	(0.028)	(0.441)	(0.402)	(0.579)
Control Variables:												
Entry Cohort Fixed Effects		+	+		+	+		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+		+	+		+	+
County Indicators			+			+			+			+
Parental Variables (for Mother and Father):												
Highest Completed Education and Income												
Highest Completed Education in STEM Field		+	+		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+			+
Mean of Older "Sibling" *MathScience*				0.269						0.274		
Mean of Younger "Sibling" *MathScience*				0.183						0.198		
Number of Randomly Matched Sibling Pairs				1,678						1,140		

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively. Only children who are enrolling in high school in 1984-91, graduating within three years, and are 18-22 years old at graduation are given a random number. First, they are matched based on mother's education, income and STEM field indicator and the rank of their random number within the 1984-87 and 1988-91 cohorts Just as for our main sample, we condition on placebo siblings having an age difference of mostly four years. Columns (1)-(6) display results when only children are matched across high schools, whereas columns (7)-(12) display results when they are matched with only children attending the same high school.

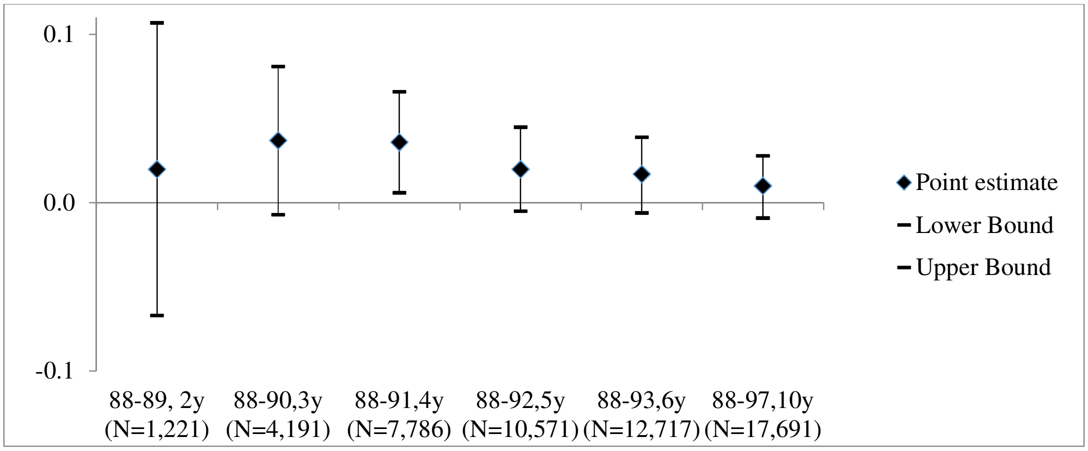
Table B12. Estimates of Spillover Effects: Heterogeneity by Parental Education and Gender

	Parameter Estimates (Standard Error)											
	Parent and Older Daughter						**Parent and Older Son**					
	STEM = 0			**STEM = 1**			**STEM = 0**			**STEM = 1**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
	First-stage: Older Sibling *MathScience*											
PilotIntro	0.040 *	0.039 *	0.043 *	0.053	0.055	0.060	0.114 ***	0.112 ***	0.104 ***	0.089 *	0.076	0.085
	(0.022)	(0.022)	(0.023)	(0.043)	(0.043)	(0.045)	(0.032)	(0.033)	(0.034)	(0.052)	(0.052)	(0.055)
Relative to overall first-stage	0.588	0.565	0.623	0.779	0.797	0.870	1.676	1.623	1.507	1.309	1.101	1.232
	Reduced-form: Younger Sibling *MathScience*											
PilotIntro	0.022	0.014	0.012	0.027	0.018	0.032	0.039	0.038	0.022	0.073 *	0.061	0.034
	(0.022)	(0.021)	(0.022)	(0.042)	(0.041)	(0.042)	(0.025)	(0.024)	(0.024)	(0.044)	(0.042)	(0.044)
	Outcome: Younger Sibling *MathScience*											
Older Sibling *MathScience*	0.564	0.368	0.279	0.518	0.334	0.528	0.344	0.336	0.215	0.828	0.801	0.404
	(0.617)	(0.562)	(0.514)	(0.856)	(0.786)	(0.778)	(0.229)	(0.218)	(0.238)	(0.611)	(0.681)	(0.520)
Control Variables:												
Entry Cohort Fixed Effects		+	+		+	+		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+		+	+		+	+
County Indicators			+			+			+			+
Parental variables (for mother and father):												
Highest Completed Education and Income		+	+		+	+		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+		+	+		+	+
HS Mean og Highest Completed Education and Income			+			+			+			+
Mean of Older Sibling *MathScience*		0.166			0.254			0.458			0.541	
Mean of Younger Sibling *MathScience*		0.165			0.246			0.177			0.231	
Number of Sibling Pairs		**3,041**			**1,072**			**2,561**			**1,112**	

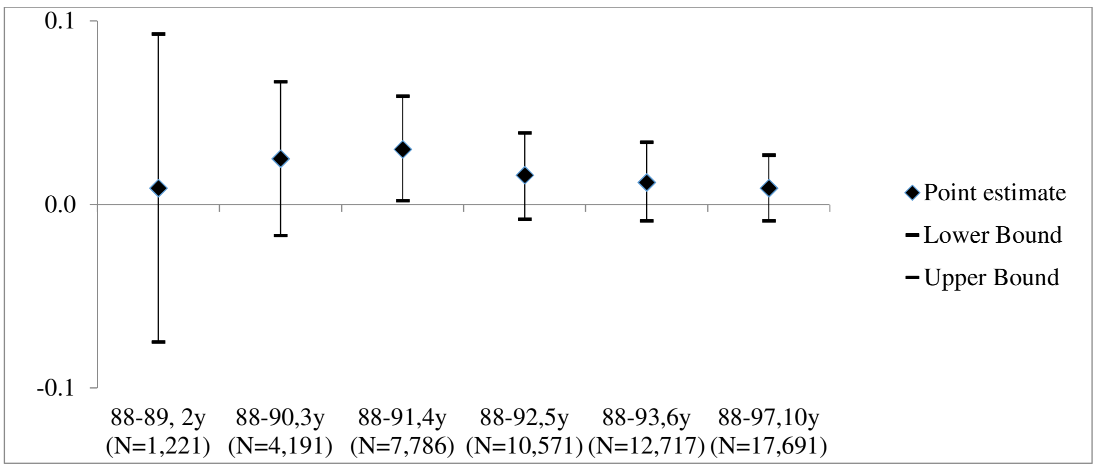
Note: The definition of STEM fields follows the definition by the National Science Foundation (NSF), which includes math, engineering, natural and technical sciences as well as some social sciences and life sciences. Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Figure B1. Reduced-form Estimates of Pilot Spillover Effects: Sibling Spacing

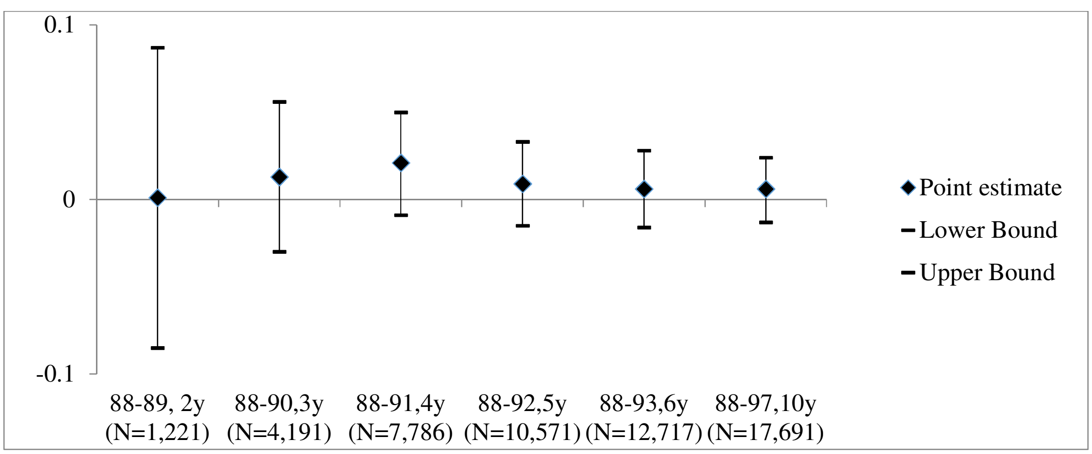
(1)



(2)

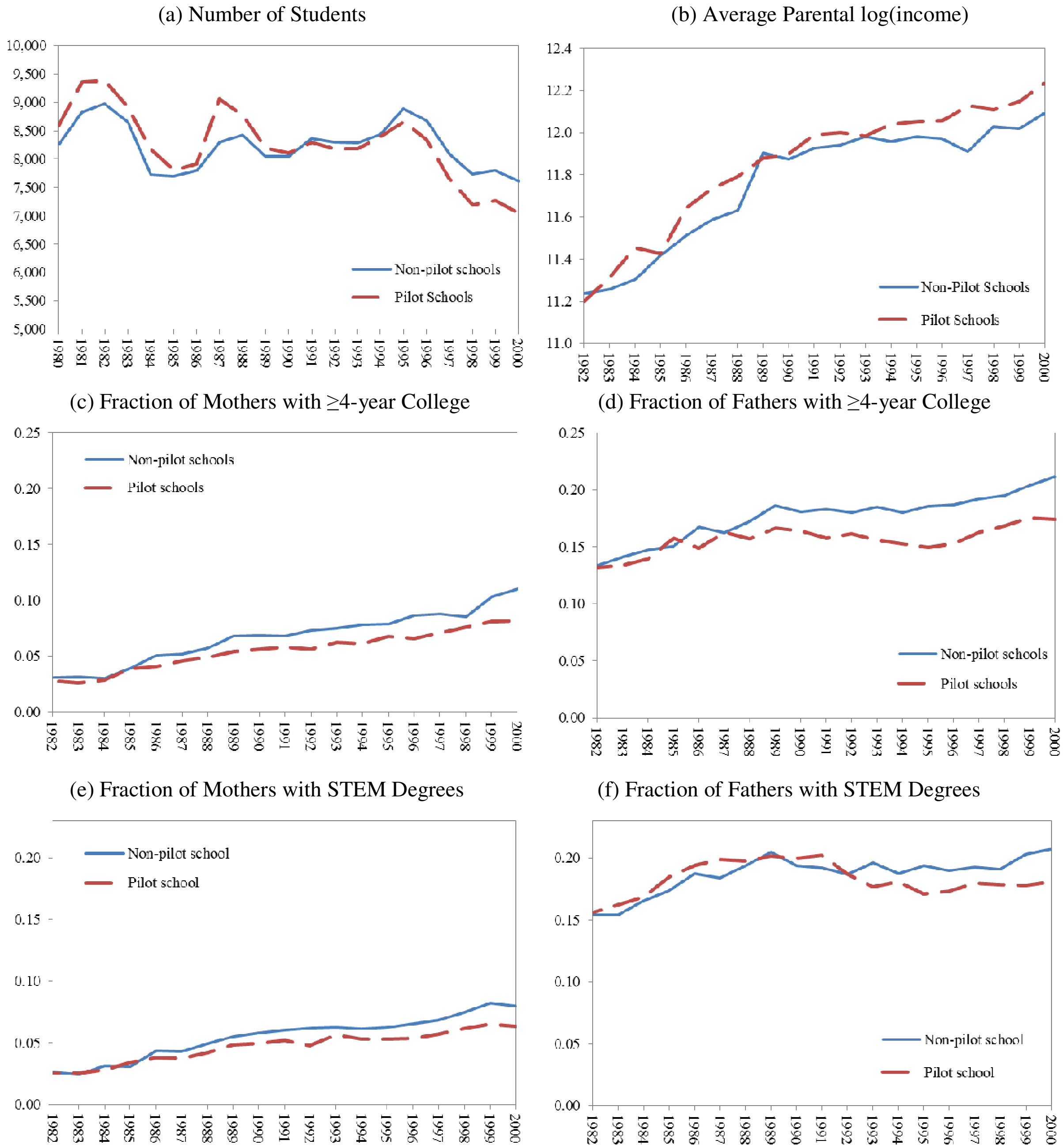


(3)



Note: Figure illustrates reduced-form coefficient estimates and 95%-confidence intervals for estimates of pilot spillover effects on younger sibling math-science choice for sibling pairs ≤ 2, ≤ 3, ≤ 4, ≤ 5, ≤ 6 and ≤ 10 years apart. (1): no control variables apart from post pilot school; (2): sibling pair gender indicators, cohort, and parental variables; (3) = (2) + county fixed effects and high school variables.

Figure B2. School Characteristics, by School Pilot Status.



Note: This Figure supplements Figure 1 by showing characteristics of the student body and their parents by high school entry cohort (on the x-axis) and by pilot status. The red dashed lines refer to the 64 schools with pilot status at any time during 1984-87, while the blue solid lines refer to the schools which never implemented the pilot scheme. Panel (a) displays the number of high school graduates within three years, panel (b) displays log total parental income, while panels (c)-(f) display the fraction of mothers and fathers with at least a 4-year college degree and a STEM degree, respectivley. Parental background variables are only sparsely available before 1982.

Appendix C. Birth-Order Effects

In Table C1 we investigate heterogeneous effects by birth order and size of the sibship. We find that first-born older siblings respond less strongly to the introduction of the pilot program *and* influence their younger sibling more strongly than later-born older siblings. For pairs where the young sibling is also the last-born in the sibship, the old sibling responds more strongly to the pilot program *and* influences the younger sibling more than in other sibships. This latter finding is not entirely consistent with the predictions from the psychological literature, that middle children compete more with first borns, while last borns do not. Our results reveal no difference between families with two siblings versus three or more siblings.

There are two competing hypothesis why first-borns influence younger siblings more than later-borns: either they exert a stronger *direct* influence on all of the younger siblings, or they exert a stronger effect due to *indirect* snowball effects. In our set-up we analyze 1,842 sibships of three ("triplets") to identify the birth order effect without imposing additional parameter restrictions. Below we provide details on how we exploit the unique features of the institutional setup and the timing of the policy changes, which implies that some "triplets" have one and some have two older siblings exogenously affected by the pilot scheme. This allows us to separately identify both the direct effect of the first *and* of the second on the third sibling, without imposing additional parameter restrictions. We are thus able to estimate birth-order effects under more general conditions than the social multiplier conditions provided by Glaeser, Sacerdote and Scheinkman (2003) and the snowball effect conditions provided by Dahl, Løken and Mogstad (2014).

Table C1. Estimates of Spillover Effects: Heterogeneity by Birth Order and Sibship Size

	Parameter Estimates (Standard Errors)																	
	Older Sibling						Younger Sibling						Number of Siblings in Sibship					
	Not Oldest			Oldest			Not Youngest			Youngest			Nsibs = 2			Nsibs > 2		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**	**(17)**	**(18)**
First-stage: Older Sibling *MathScience*																		
PilotIntro	0.147 ***	0.150 ***	0.150 ***	0.048 **	0.047 **	0.048 **	0.032	0.035	0.035	0.085 ***	0.085 ***	0.081 ***	0.069 ***	0.060 ***	0.060 **	0.066 ***	0.077 ***	0.078 ***
	(0.040)	(0.038)	(0.038)	(0.020)	(0.018)	(0.019)	(0.031)	(0.030)	(0.030)	(0.021)	(0.020)	(0.021)	(0.025)	(0.023)	(0.024)	(0.025)	(0.024)	(0.025)
Relative to overall first-stage	2.162	2.174	2.174	0.706	0.681	0.696	0.471	0.507	0.507	1.250	1.232	1.174	1.015	0.870	0.870	0.971	1.116	1.130
Reduced-form: Younger Sibling *MathScience*																		
PilotIntro	0.016	0.015	0.01	0.041 **	0.034 **	0.025	0.031	0.030	0.000	0.038 **	0.030 *	0.030 *	0.033	0.023	0.023	0.038 *	0.035 *	0.017
	(0.032)	(0.031)	(0.032)	(0.016)	(0.016)	(0.016)	(0.027)	(0.026)	(0.026)	(0.017)	(0.017)	(0.017)	(0.021)	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)
Outcome: Younger Sibling *MathScience*																		
Older Sibling *MathScience*	0.107	0.097	0.070	0.839 *	0.730 *	0.508	0.988	0.859	-0.012	0.449 **	0.357 *	0.367 *	0.473	0.391	0.387	0.577	0.449	0.215
	(0.220)	(0.204)	(0.231)	(0.434)	(0.406)	(0.367)	(1.188)	(0.960)	(0.657)	(0.221)	(0.203)	(0.220)	(0.318)	(0.341)	(0.355)	(0.363)	(0.279)	(0.266)
Control Variables:																		
Entry Cohort Fixed Effects		+	+		+	+		+	+		+	+		+	+		+	+
Sibling Pair Gender Composition		+	+		+	+		+	+		+	+		+	+		+	+
County Indicators			+			+			+			+			+			+
Parental Variables (for Mother and Father):																		
Highest Completed Education and Income		+	+		+	+		+	+		+	+		+	+		+	+
Highest Completed Education in STEM Field		+	+		+	+		+	+		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+			+			+			+
Mean of Older Sibling *MathScience*		0.295			0.337			0.325			0.329			0.343			0.314	
Mean of Younger Sibling *MathScience*		0.177			0.193			0.198			0.186			0.190			0.189	
Number of Sibling Pairs		**1,634**			**6,152**			**2,489**			**5,297**			**3,764**			**4,022**	

Note: Standard errors are clustered by older siblings. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

To illustrate, we simplify the linear projections (3) and (4) to exclude other characteristics (subscript *f* suppressed):

$$MathScience_{old} = \pi_0 + \gamma PilotIntro_{old}$$

$$MathScience_{young} = \beta_0 + \beta_1 MathScience_{old}$$

Assume there are three siblings, where the older siblings affect younger siblings. If only the oldest was affected by the pilot scheme, then we have the system of equations:

$$MathScience_1 = \pi_0 + \gamma PilotIntro_1$$

$$MathScience_2 = \beta_{02} + \beta_{12} MathScience_1$$

$$MathScience_3 = \beta_{03} + \beta_{13} MathScience_1 + \beta_{23} MathScience_2$$

where 1 denotes the oldest sibling, 2 the middle, and 3 the youngest. Using exogenous variation from the cost of obtaining Math-Science for the oldest sibling, we could identify the reduced form parameters:

$$\frac{\partial MathScience_2}{\partial PilotIntro_1} = \beta_{12}\gamma = \delta_2$$

$$\frac{\partial MathScience_3}{\partial PilotIntro_1} = (\beta_{23}\beta_{12} + \beta_{13})\gamma = \delta_3$$

The total peer effect on the Math-Science choice of the younger sibling *j* is given by δ_j divided by the first-stage coefficient γ. Comparing the estimated total effects across younger siblings in this setting, we could follow Dahl, Løken and Mogstad (2014) and assume that the direct effect of an older sibling on all younger siblings is identical: $\beta_{12} = \beta_{13} = \beta_1$ and $\beta_{23} = \beta_2$, and then identify snowball effects: $\beta_2\beta_1$. The second sibling identifies the direct effect, β_1, as δ_2 divided by γ. Subtracting off this direct effect, the snowball effect on the third sibling, $\beta_2\beta_1$, is given by the difference $\delta_3 - \delta_2$ divided by γ.[30] This is, however, an unreasonable assumption in our setting as the peer effect varies with sibling age difference. Alternatively Glaeser et al. (2003) identify social multipliers under the assumption that $\beta_{12} = \beta_{23} = \mu$ and $\beta_{13} = \mu\rho$ meaning that the effect on the immediate younger sibling is the largest and identical independent of birth order. This is, however, also an unreasonable assumption in our

[30] More generally, they estimate a decay function to allow the direct effects to decay over time and assume the functional form of the decay function is the same for the same distance.

setting as age difference and sibling interactions (in terms of birth order and gender composition) matter for the strength of the peer effect.

We exploit that we have variation in how many older siblings were exposed to the pilot scheme in order to avoid having to make unreasonable parametric restrictions. If the two oldest siblings were affected by the pilot scheme, then:

$$MathScience_1 = \pi_0 + \gamma PilotIntro$$

$$MathScience_2 = \beta_{02} + \beta_{12} MathScience_1 + \gamma PilotIntro$$

$$MathScience_3 = \beta_{03} + \beta_{13} MathScience_1 + \beta_{23} MathScience_2$$

$$= \beta_{03} + \beta_{13}(\pi_0 + \gamma PilotIntro) + \beta_{23}(\beta_{02} + \beta_{12}(\pi_0 + \gamma PilotIntro) + \gamma PilotIntro)$$

Exploiting the exogenous variation that some have one and some have two older siblings affected by the pilot scheme, we can separately identify both the direct effect of the first *and* of the second on the third sibling, without imposing additional parameter restrictions. Again, looking at the reduced form parameters:

$$\frac{\partial MathScience_2}{\partial PilotIntro} = (\beta_{12} + 1)\gamma = \delta_2 + \gamma = \theta_2$$

$$\frac{\partial MathScience_3}{\partial PilotIntro} = (\beta_{23}\beta_{12} + \beta_{23} + \beta_{13})\gamma = \delta_3 + \beta_{23}\gamma = \theta_3$$

First, the reduced form for the second sibling identifies the direct effect of the first on the second sibling, β_{12}, as δ_2 divided by γ (or alternatively as θ_2 divided by γ minus one: $\beta_{12} = \frac{\delta_2}{\gamma} = \frac{\theta_2}{\gamma} - 1$). Second, we identify the direct effect of the second sibling on the third sibling, β_{23}, as $\theta_3 - \delta_3$ divided by : $\beta_{23} = \frac{\theta_3 - \delta_3}{\gamma}$. Third, we identify the direct effect of the first on the third sibling, β_{13}, as δ_3 divided by γ minus $\beta_{23}\beta_{12}$. That is, $\beta_{13} = \frac{\delta_3}{\gamma} - \beta_{23}\beta_{12}$.

We can therefore identify the direct *and* indirect (multiplicative) average effects between all three siblings without additional parameter restrictions.

To this end we exploit that we observe 1,842 triplets; 564 for which the two oldest siblings entered high school during the pilot period. The catch here is that we only observe two triplets for which both older siblings were unexpectly exposed to the pilot introduction, 90 where only the oldest, and 56 where only the second was exposed. We therefore lump together all the pilot schools for this part of the analysis. We thus disregard whether the pilot was unexpectedly introduced or not which may entail a bias as some older siblings may have chosen high school based on the availability of the pilot. However, Joensen and Nielsen (2009) find that the high school choice is neither sensitive to distance to school nor its interaction with pilot status, and they find similar causal effects on earnings when not distinguishing between students who were unexpectedly exposed to the pilot and those who enrolled after the pilot status was announced. Our estimates of the direct effects are: β_{12} = 0.24, β_{13} = -0.23, β_{23} = -0.42, but too noisy to draw strong inference on direct and indirect birth-order effects. These estimates suggest a strong positive direct effect of the first- on the second-born, but negative direct effects on the third-born. The pattern is consistent with Adams (1972) who suggests that second- and middle-born are more likely to conform to the first-born child, while the last-born is more likely to be a non-conformer.