

Peer Effects in Math and Science*

Juanna Schrøter Joensen
Stockholm School of Economics

Helena Skyt Nielsen
Aarhus University

February 3, 2015

Abstract:

In this paper, we specify and test a model of how the aggregate skill supply is shaped by institutions and social peer groups. First, we exploit a universal policy reform signaling less importance of advanced math-science in high school. We show how it amplified the fall in skill supply in peer groups with a stronger math-science norm. We also document a gender convergence in math-science skills. This is mainly triggered by boys strongly crowding in the policy reform as they face a higher social cost of not conforming. Second, we analyze the underlying mechanism by estimating causal sibling peer effects in the math-science choice. We exploit quasi-experimental variation stemming from a pre-reform pilot scheme. The pilot induced some older siblings to choose advanced math-science at a lower cost, while not directly affecting the course choices of younger siblings. Therefore, any influences of the pilot scheme on the younger siblings may be attributed to the peer influence of the older sibling. Our results suggest that peer effects among siblings are strongest among closely spaced siblings, in particular brothers. We argue that competition is likely the driving force behind younger siblings conforming to their older siblings' choices.

JEL Classification: I21, I24, J24, J16, J12, Z13.

Keywords: Social interaction, extrinsic and intrinsic incentives, siblings, high school curriculum, skill formation.

* Contact details: Joensen, Department of Economics, Stockholm School of Economics, email: Juanna.Joensen@hhs.se. Nielsen: Department of Economics and Business, Aarhus University, email: hnielsen@econ.au.dk. We appreciate enlightening discussions with and comments from Anders Björklund, Tore Ellingsen, Ina Ganguli, Leonie Gerhards, Maria Knoth Humlum, Eric Maurin, Magne Mogstad, Julia Nafziger, Torsten Persson, Barbara Petrongolo, and participants at SOLE 2014, ELE 2014, IWAAE 2013 and seminar participants at the CEP-LSE Education Group, Copenhagen Business School, Federal Reserve Bank of New York, Institute for International Economic Studies, Lancaster University, Oslo University, Stavanger University, SOFI Stockholm University, Sussex University, University of Southern Denmark, York University, and Aarhus University. We thank 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. Peer 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. As a consequence, social interaction effects may have contributed to increased inequality and the sluggish skill development in terms of slowing down both college preparedness and college enrollment (Goldin and Katz, 2008).

In this paper, we focus on social interactions within the family. We contribute to the emerging research trying to unravel the role of the family in skill formation and human development more generally.¹ We document that sibling spillovers are possibly large and we provide a framework for understanding the plausible mechanisms. 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 in schooling have recently been examined.² In this paper, we recognize that family background is multifaceted and we go beyond simple sibling correlations and estimate causal effects in order to understand the large and widespread consequences of the complex family component that siblings share.

We focus on the role of institutions in skill formation and contribute to the scarce literature on the nature of social interaction effects in education. In their seminal paper, Akerlof and Kranton (2002) provide a review of the sociology of education literature and incorporate the ideas in an economic framework. They highlight that a main motivation for educational investments is the identity utility gain from fitting into a social group. Empirical evidence of the relevance of these mechanisms is scarce, but pivotal for better understanding *how* economic shocks and policy interventions influence educational decisions.

The main contribution of this paper is twofold. First, we specify and test a model of how aggregate skill supply is shaped by institutions and their interactions with the social environment in which individuals make their educational decisions. Second, we investigate the underlying mechanisms by estimating causal spillover effects at the sibling pair level.

We set up a model of how individual educational choices are determined by not only *extrinsic* incentives (material costs and benefits), but also *intrinsic* incentives (social norms or identity) and their

¹ 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.

² See e.g. Solon (1999) and Black and Devereux (2010) for reviews on intergenerational correlations and Mazumder (2008), Björklund and Salvanes (2010) and Björklund and Jäntti (2012) for sibling correlations.

interaction. This model builds on the theoretical framework of two recent papers by Benabou and Tirole (2011) and Jia and Persson (2014) where complementarity is not imposed, but conformity arises endogenously in equilibrium as individuals infer the social norm from their peers' educational choices.³ Extrinsic incentives depend on the monetary payoff as well as prizes, laws, or curricula which not only impact costs of educational choices directly, but also send signals about the values of schools and society. Intrinsic incentives may depend on individuals' social environment; primarily, siblings and parents, but also schoolmates. Younger siblings get a utility gain (loss) of (not) conforming to their older sibling and the identity utility gain depends on how many conformers are in their peer group. We derive two testable predictions of how responses to changes in extrinsic incentives (in our case school curricula) vary based on their interaction with (i) the strength of social norms and (ii) intrinsic incentives. We test these two predictions on Danish administrative data, where individuals are linked with their siblings, parents, schoolmates, as well as their schoolmates' siblings and parents. We focus on the choice of advanced coursework in mathematics and science in high school, which is a prerequisite for increasing the supply of college graduates in science, technology, engineering, and mathematics (STEM). The major high school reform in Denmark in 1988 loosened up the bundling of advanced math and science courses, and the number of students opting for these advanced skills dropped significantly. We exploit this reform as an exogenous decrease in extrinsic incentives and test how choices differentially respond to the universal policy change depending on the strength of the STEM norm in the peer group and their intrinsic incentives of conforming to their sibling's choice. We find strong empirical support for both model predictions. First, we find that extrinsic incentives are more strongly *crowded in* for younger brothers and sibling pairs where the older sibling chose advanced math and science. This means that a curriculum change – signaling school values and lowering the costs of opting for math and science – will cause larger changes in behavior in environments with a stronger STEM norm and will also be more strongly amplified among boys. Second, we find that boys have a higher intrinsic cost of non-conforming to an older sibling who has chosen math and science. This could indicate that there is more social pressure on boys to choose math and science in the baseline, so when the choice set is loosened up and schools signal the lesser importance of this combination, it is more strongly crowded in by boys as the social-to-material payoff is higher for the marginal boy than the marginal girl.⁴

³ In this respect, our model extends Akerlof and Kranton (2002).

⁴ This is also consistent with Joensen and Nielsen (2014) who find that the baseline monetary payoff for the marginal girl to choose advanced math and science (when only 20% of girls in the high school cohort are choosing it) is substantial,

Having established how educational choice responses to policy changes systematically interact with social norms and intrinsic incentives, we turn to better understanding the underlying mechanisms by estimating causal sibling spillover effects. Estimating the causal effect of social interactions is challenging due to simultaneity, correlated unobservables, and endogenous peer group membership (Manski, 1993; 1995). We study naturally occurring peer groups and exploit exogenous variation in the cost of taking up advanced math and science in high school among a partial population (Moffitt, 2001).⁵ We exploit the fact that some older siblings in Denmark in 1984-1987 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; 2014) we employed this pilot scheme to investigate the impact on the individuals' own educational choices, subsequent careers, and earnings. Those unexpectedly exposed to the pilot scheme got more advanced college degrees and substantially higher earnings. In this paper, we analyze whether there were spillover effects on younger siblings who were unexposed themselves. Any influence of this pilot scheme on younger siblings' course choices can be interpreted as a causal peer effect, since the pilot scheme only directly reduced the cost of choosing advanced math and science for older siblings. The exogenous increase in extrinsic incentives for the older sibling is thus essentially an increase in intrinsic incentives for the younger sibling in the presence of identity concerns and utility costs of not conforming. We find that younger siblings are 3.5 percentage points more likely to choose math and science if their older sibling was exposed to the pilot scheme. Since the first-stage estimate is 7 percentage points, this implies a peer influence of older siblings on younger siblings of about 0.5. This means that by affecting the choice of the older sibling, about half of the effect is expected to spill over on the younger sibling's choice. More generally, this suggests that knowledge about the social peer group is important to predict the total impact of education policies, and that policies targeted at influential peers (such as older siblings) would have more widespread effects.

There are many reasons why siblings are influential peers. Siblings are the first peers one closely interacts with and for most parts they entail a lifelong relationship. Therefore, peer effects from close social interaction between siblings may be extremely important (Buhrmester, 1992). Rigorous economics research on the importance of social interactions among siblings is scarce. Butcher and Case (1994) find that the level of education among females decreases with the presence of any sister

while it is lower (approaching zero) for the marginal boy (when more than 50% of boys choose advanced math and science).

⁵ Our study is thus methodologically similar to the study of social interaction effects in program participation by Dahl, Løken and Mogstad (2014), Avvisati, Gurgand, Guyon and Maurin (2014), and Dahl, Kostøl and Mogstad (2014).

in the sibship, and that this effect gradually vanishes for more recently born cohorts. They argue that the presence of a second daughter in the household changes the reference group of the first daughter. Qureshi (2013) and Adermon (2013) study education of sibling pairs in Pakistan and Sweden, respectively. Qureshi (2013) 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. Adermon (2013) finds no spillover effects among siblings from extending compulsory schooling laws in Sweden. Sibling spillover effects have also been documented in parental leave taking among brothers (Dahl, Mogstad and Løken, 2014), from newborn health to educational outcomes (Breining, Daysal, Simonsen and Trandafir, 2015), and adolescent smoking, drinking and marijuana use (Altonji, Cattan and Ware, 2013).

We find strong positive correlations between math and science choices of siblings. Our results suggest that there is substantial heterogeneity in peer effects – both in terms of how strongly the older sibling responds to extrinsic incentives and how their choice spills over on the younger sibling through intrinsic incentives. The causal peer effects persist among closely spaced siblings, and the significance depends on the gender composition of the sibling pair. We find largest and most significant peer effects for relatively closely spaced brothers. First-born siblings are the most influential peers and parental education is also important for spillovers. We provide 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.

The remainder of the paper unfolds as follows: Section 2 specifies and tests the predictions of an economic model of educational choices and peer effects. Section 3 discusses identification of social interaction effects and presents the institutional background which our empirical strategy relies on. Section 4 describes the data, while section 5 presents the empirical analysis of social interaction effects in the choice of math and science in high school. Section 6 investigates mechanisms and heterogeneity in peer effects. Section 7 concludes the paper.

2. A Framework Linking Educational Choices and Peer Effects

In this section we lay out an economic model of how an individual makes educational choices conditional on the institutions and social environment. We derive two model predictions and test them empirically based on differential course choice responses to a major institutional reform across peer groups with different social norms and individuals facing different intrinsic incentives.

Our model describes how an individual’s choice can be determined by not only extrinsic incentives, but also intrinsic incentives and their interaction. Extrinsic incentives include monetary pay but also laws, prizes, and curricula which signal school values. Intrinsic incentives depend on the social environment. In this section we focus on how intrinsic incentives vary with older sibling’s choice, while we try to unravel heterogeneity in this intrinsic cost due to differences in the strength of ties and sibling interactions (e.g. competition and sibling rivalry) in Section 6. The model builds on Benabou and Tirole (2011) and Jia and Persson (2014). It illustrates how we can explain the falling trend in the number of individuals taking advanced math and science courses, and understand the role of the social environment in explaining this trend. Primarily, we analyze the first-order peer effect of how younger siblings’ intrinsic motivation to opt for math and science depends on older siblings’ choice of the same course combination. More broadly, we analyze how choices and incentive effects are affected by the social environment through other peers; including parents, schoolmates, and schoolmates’ older siblings and parents.

Figure 1. Fraction of High School Cohort with advanced Math and Science.

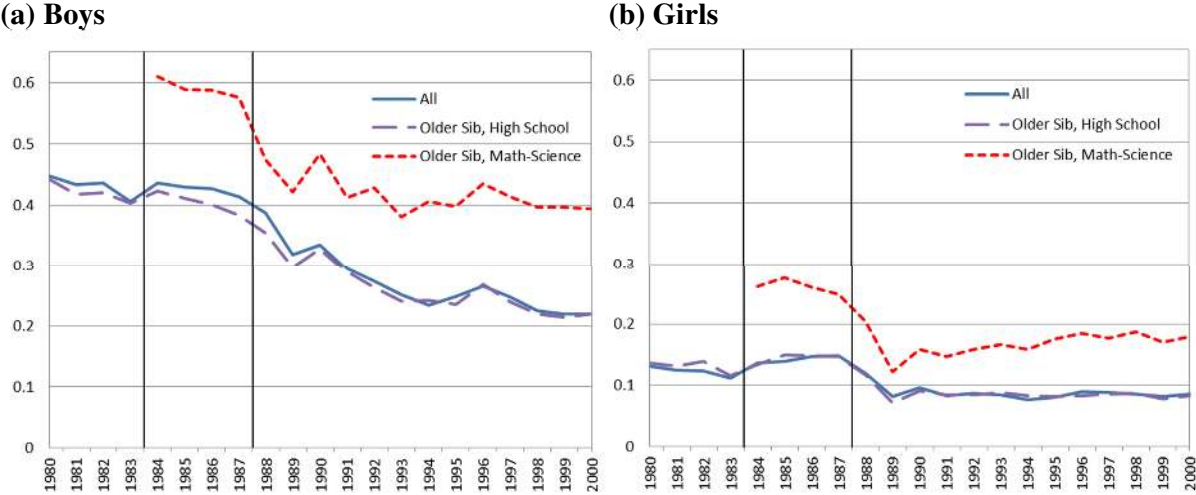


Figure 1 shows the fall in the fraction of the high school cohort who acquired advanced math and science skills over the period 1980-2000 in Denmark. From the figure it is evident that two factors appear important: (i) older siblings’ course choices and (ii) school curricula. Individuals with an older sibling who chose math and science are more likely to also choose math and science. During the pilot period 1984-87 (explained in detail in Sections 3.2 and 3.3) the falling trend leveled off as an additional advanced math and science course package was introduced at pilot schools. The major reform in 1988 that broke up the course bundling and gave students (as good as) free course choice caused a major

drop in the number of students opting for advanced math and science.⁶ The proportion of girls with advanced math and science stabilized around 8% after 1988, while the gradual decline continued for boys such that only half as many boys chose math and science in 2000 compared to 1980. The gender gap in these advanced skills thus fell by 57% over this 20-year period: from a 31 percentage point difference in 1980 to a 13 percentage point difference in 2000.

We specify a model that highlights how this development can be explained by the interaction between extrinsic incentives and social multipliers across peer groups. We derive this model prediction in Section 2.1 and test it in Section 2.3 below. We show that the time patterns in Figure 1 can be explained by a stronger social multiplier for boys and for younger siblings with an older Math-Science sibling. A model without intrinsic motivation and social concerns would not be able to explain these time patterns without gender-biased changes in preferences or relative skill prices. In our framework, these predictions arise from simply allowing younger sibling's course choice to depend on the institutional incentives and the social environment they face. We distinguish between two cases: (a) the older sibling does not choose Math-Science and (b) the older sibling chooses Math-Science.

(a) **Non-Math-Science older sibling** ($MathScience_{old} = 0$)

If the older sibling does not choose Math-Science, then the younger sibling's utility of choosing Math-Science is given by:

$$U_0 = (b - e_{old}^0 - \varepsilon) MathScience_{young} + \mu E [\varepsilon \mid MathScience_{young}] \quad (1)$$

where b denotes extrinsic incentives, which may vary across schools and time; e.g. due to variation in monetary pay, college access, or curriculum design. The term $e_{old}^j + \varepsilon$ denotes intrinsic costs of deviating from the choice of the older sibling. The component $e_{old}^j \equiv e(MathScience_{old})$ for $MathScience_{old} = j \in \{0,1\}$ depends on the older sibling's course choice whereas the component, ε , is unrelated to the older sibling's particular course choice. e_{old}^j is the average intrinsic cost for younger siblings who choose differently than their older sibling. This intrinsic cost is the same for everyone in the same peer group, but may vary across peer groups. The primary source of heterogeneity is ε , which captures deviations from the average intrinsic cost. $E[\varepsilon \mid MathScience_{young}]$ is the truncated mean over ε for everyone in the peer group making the same Math-Science choice, and it determines the costs to the *identity* of the relevant course choice.⁷ We define the gain in identity by *conforming* with the older

⁶ The reform may be viewed as a movement towards a "Shopping Mall High School" as first described by Powell, Farrar and Cohen (1985) and discussed by Akerlof and Kranton (2002). This reform is explained in detail in Section 3.4.

⁷ Note we choose to use the term identity throughout, but this social factor can be thought of more broadly as a social norm, social reputation, social esteem, or self-image.

sibling's choice to be $\Delta = E[\varepsilon \mid \text{MathScience}_{young} = 0] - E[\varepsilon \mid \text{MathScience}_{young} = 1]$, where the first term is the *pride* of conforming with the older sibling and the second term is the *prejudice* of choosing a Math-Science identity not conforming with the older sibling.⁸ μ is the weight placed on identity, which may vary with the strength with which social norms are held. Particularly, the weight on identity varies across different peer groups. We define the relevant peer group to be the students in the same cohort at the same school.⁹

The model implies that the younger sibling chooses $\text{MathScience}_{young} = 1$ if and only if $\varepsilon < \varepsilon_0^*(b, e_{old}^0, \mu)$, where $\varepsilon_0^*(b, e_{old}^0, \mu)$ defines the cut-off below which younger siblings choose advanced Math-Science as a function of extrinsic and intrinsic incentives, and the weight placed on identity. Optimizing younger siblings thus balance the net extrinsic and intrinsic motivation of Math-Science with the net gains in identity of choosing the same as their older sibling:

$$b - e_{old}^0 - \varepsilon_0^*(b, e_{old}^0, \mu) = \mu \Delta(\varepsilon_0^*) \quad (2)$$

where $\Delta(\varepsilon_0^*) = E[\varepsilon \mid \varepsilon > \varepsilon_0^*] - E[\varepsilon \mid \varepsilon < \varepsilon_0^*]$ is the equilibrium gain in identity. We assume that $E[\varepsilon] = 0$, which means that $\Delta(\varepsilon_0^*)$ is positive by definition as the first pride-term is positive and the second prejudice-term is negative.¹⁰ The fraction of sibling pairs where the younger sibling chooses Math-Science and does not conform with his or her older sibling is given by $F(\varepsilon_0^*(b, e_{old}^0, \mu))$, as those with $\varepsilon < \varepsilon_0^*(b, e_{old}^0, \mu)$ choose Math-Science. We label these non-conformers; see Figure A1.

The sign of the derivative of conformity, $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon}$, is crucial for the properties of equilibrium and comparative statics. If the cut-off, ε^* , increases so more younger siblings choose Math-Science, then both the pride and the prejudice term increase. We can determine which one increases more (i.e. sign the derivative of conformity) under the following mild distributional assumptions: continuously differentiable cumulative distribution function (CDF), $F(\varepsilon)$, and a symmetric unimodal probability density function (pdf), $f(\varepsilon)$. This implies that $\Delta(\varepsilon)$ has a unique interior minimum at $\varepsilon = 0$.¹¹ Thus $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon} < 0$ for low values of ε^* , while $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon} > 0$ for high values of ε^* . This means that Math-Science

⁸ We choose the terms pride (prejudice) throughout, but they can be interpreted more broadly as honor (stigma) or virtue (guilt). These terms were also adopted by Ellingsen and Johannesson (2008) who provide a related model where some agents have identity concerns and its value depends on the principal. In our model, identity concerns instead depend on the educational choices of the peer group in equilibrium.

⁹ Alternatively, we define the relevant peer group to be the students of the same gender in the same school-cohort, as social comparisons within classrooms tend to be made within gender (Jalava et al., 2014).

¹⁰ Over the whole peer group, however, the gain in identity is a zero-sum game. Since younger siblings do not know ε before they make course choices, the ex-ante expected value is zero, $E[\varepsilon \mid \text{MathScience}_{young}] = 0$. This is seen by integrating over all individuals in equilibrium: $\int_{-\infty}^{\varepsilon_0^*} E[\varepsilon \mid \varepsilon < \varepsilon_0^*] d\varepsilon + \int_{\varepsilon_0^*}^{\infty} E[\varepsilon \mid \varepsilon > \varepsilon_0^*] d\varepsilon = 0$.

¹¹ See Jewitt (2004) for a proof and Figure A2 in Appendix A for an illustration.

choices for siblings are strategic complements (i.e. reinforcing one another) when few choose Math-Science as the prejudice increases by more than the pride. Therefore, the gain in identity from conforming with the older non-Math-Science sibling decreases and more younger siblings become non-conformers by choosing to opt for Math-Science. Conversely, choices are strategic substitutes (i.e. offsetting one another) when many choose Math-Science. The reputational return is lowest when choosing Math-Science is the modal choice. This is illustrated in Figure A2 in Appendix A.

(b) **Math-Science older sibling** ($MathScience_{old} = 1$)

If the older sibling chooses Math-Science then the younger sibling's utility is given by:

$$U_1 = MathScience_{young}b - (1 - MathScience_{young})(e_{old}^1 + \varepsilon) - \mu E[\varepsilon | MathScience_{young}] \quad (3)$$

where $e_{old}^1 + \varepsilon$ denotes the intrinsic cost of choosing differently than the older sibling. We assume that all other parameters and the distribution of the ex-ante unknown intrinsic cost, ε , are the same for all younger siblings. Symmetry of the distribution of ε means that the younger sibling chooses $MathScience_{young} = 1$ if and only if $\varepsilon < \varepsilon_1^*(b, e_{old}^1, \mu)$, where $\varepsilon_1^*(b, e_{old}^1, \mu)$ denotes the cut-off value below which the younger sibling chooses Math-Science – just like their older sibling. Again the younger sibling balances net gains such that:

$$b + e_{old}^1 + \mu \Delta(\varepsilon_1^*) = \varepsilon_1^*(b, e_{old}^1, \mu). \quad (4)$$

where $\Delta(\varepsilon_1^*) = E[\varepsilon | \varepsilon < \varepsilon_1^*] - E[\varepsilon | \varepsilon > \varepsilon_1^*]$ is defined as equilibrium conformity to an older Math-Science sibling, and analogously to before, the first term is the pride of choosing Math-Science like the older sibling and the second term is the prejudice of not conforming. The fraction of sibling pairs where both the younger and older sibling choose Math-Science is thus given by $F(\varepsilon_1^*(b, e_{old}^1, \mu))$. We label these conformers.

From the decision rules (2) and (4) emerges the immediate prediction that more younger siblings choose Math-Science when their older siblings also chose Math-Science. As the cut-off is higher when the older sibling has Math-Science $\varepsilon_1^*(b, e_{old}^1, \mu) > \varepsilon_0^*(b, e_{old}^0, \mu)$ it also implies that more younger siblings choose Math-Science conforming with their older siblings, $F(\varepsilon_1^*(b, e_{old}^1, \mu)) > F(\varepsilon_0^*(b, e_{old}^0, \mu))$. That is, there are relatively more conformers than non-conformers among the younger siblings choosing Math-Science. This prediction is consistent with the data (presented in detail in Section 4) and Figure 1: 28% of younger siblings choose Math-Science conforming to their older siblings, while only 14% are non-conformers and choose Math-Science even if their older siblings did not. The intuition behind this result is that younger siblings receive both extrinsic benefits and additional intrinsic benefits when conforming to their older siblings. The social norm could further be

strengthened by having more Math-Science peers; for example, parents with a STEM field education or schoolmates with more Math-Science tradition in their families. This variation will also be exploited in the empirical analysis.

The gender of the older sibling seems to matter for younger sisters, as 21% conform to an older sister while only 14% conform to an older brother. 42% of younger brothers are conformers, independently of their older sibling's gender.¹² These patterns might suggest that the intrinsic costs of not conforming are higher for boys than for girls, and we will show that this might explain why the abovementioned curriculum changes have a different impact for boys than girls. As we move to the left in Figures A1-A2 and Math-Science becomes a rarer choice, the prejudice becomes even stronger if having a non-Math-Science older sibling and the pride becomes even stronger when having an older Math-Science sibling. As gender is a salient part of identity (Akerlof and Kranton, 2002; Bertrand, 2010; Bertrand et al., 2013) and Math-Science choices vary by gender (Joensen and Nielsen, 2014; Buser, Niederle and Oosterbeek, 2014) the model also predicts that the identity concerns vary by gender. To shed light on these issues and to understand how they relate to potential gender stereotypes (Steele, 1997; Babcock and Laschever, 2009) of Math and Science, we derive and test model predictions in the following three subsections.

2.1. *Effects of Extrinsic Incentives*

First, we investigate how an increase in extrinsic benefits, b , influences the fraction of younger siblings choosing Math-Science. The slope of aggregate Math-Science high school graduate supply:

$$\frac{\partial F(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial b} = f(\varepsilon_j^*(b, e_{old}^j, \mu)) \frac{1}{1 + \mu \frac{\partial \Delta(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial \varepsilon}} > 0 \quad (5)$$

implies that increasing extrinsic benefits increases the fraction of younger siblings choosing Math-Science independently of their older sibling's choice. The last term denotes the *social multiplier*.¹³ Since it is less likely that a younger sibling is a non-conformer when having a non-Math-Science older

¹² The numbers referred to may be found in Table B3 and Figure A1 in Appendix A.

¹³ Benabou and Tirole (2006) show that the equilibrium of this model is uniquely determined if the denominator of the social multiplier is positive. That is, if $1 + \mu \frac{\partial \Delta(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial \varepsilon} > 0$. This guarantees that μ is not large enough to create multiple equilibria when Math-Science choices are strategic complements.

sibling (13-14%) than when having a Math-Science older sibling (72%),¹⁴ it is the case that: $f(\varepsilon_0^*(b, e_{old}^0, \mu)) < f(\varepsilon_1^*(b, e_{old}^1, \mu))$. Moreover,

$$\frac{\partial \Delta(\varepsilon_0^*(b, e_{old}^0, \mu))}{\partial \varepsilon} < \frac{\partial \Delta(\varepsilon_1^*(b, e_{old}^1, \mu))}{\partial \varepsilon} \quad (6)$$

the marginal non-Math-Science older sibling ($MathScience_{old} = 0$) is in a region where fewer younger siblings choose Math-Science, compared to those with a Math-Science older sibling ($MathScience_{old} = 1$). This is illustrated in Figure A2 and means that the marginal non-Math-Science older sibling is a strategic complement rather than substitute. If both are strategic complements, then the complementarity is stronger for the marginal non-Math-Science older sibling. Extrinsic incentives are crowded in by identity when choices are strategic complements, $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon} < 0$, and crowded out by identity when choices are strategic substitutes, $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon} > 0$. Thus extrinsic incentives for younger siblings are more likely to crowd in (rather than crowd out) intrinsic incentives when having a non-Math-Science older sibling. However, it is ambiguous whether increasing extrinsic incentives, b , will lead to a smaller (or larger) increase in the probability of choosing Math-Science for younger siblings with a non-Math-Science older sibling as their first term in (5) is lower even if the second term is larger. Overall, when Math-Science is rarer than the modal choice for *both* groups, then the model does not deliver strong predictions on whether the fraction of conformers or non-conformers will increase more in response to stronger extrinsic incentives.

In general, the way responses to extrinsic incentives change with the social multiplier has implications for how Math-Science choices will be differentially affected across peer groups and schools depending on the predetermined choices of the older siblings.¹⁵ Even if all schools are subject to the same curriculum and experience the same curriculum change, they will not experience the same change in the fraction of students choosing Math-Science. A unique feature of our institutional setting allows us to test the prediction that as $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon}$ monotonically increases from being negative when $P(MathScience_{young} = 1)$ is small to being positive when the probability to choose Math-Science is large; i.e. the social multiplier is smaller for conformers than for non-conformers. The social multiplier is thus smaller if $MathScience_{old} = 1$ than if $MathScience_{old} = 0$. This implies that the same increase

¹⁴ More specifically, these refer to the conditional probabilities $P(MathScience_{young} = 1 | MathScience_{old} = 0)$ and $P(MathScience_{young} = 0 | MathScience_{old} = 1)$ for the non-conforming younger siblings; see Figure A1 in Appendix A. Table B3 reveals that 7% (4%) of younger sisters and 28% (21%) of younger brothers are non-conformers when having a non-Math-Science older sister (brother), while 79% (86%) of younger sisters and 58% (58%) of younger brothers are non-conformers when having a Math-Science older sister (brother).

¹⁵ This prediction extends to all peers affecting intrinsic incentives; e.g. parents, family, friends, and friends' family.

in extrinsic incentives leads to a smaller increase in Math-Science take-up of younger siblings with Math-Science older siblings. That is, when a higher fraction of younger siblings are choosing Math-Science, there is more crowding out (or less crowding in) via identity (or social reputation) mechanisms. Testing this prediction requires enough variation in the cut-off, ε_j^* , which we have because of the phase-in of the pilot-scheme in place for the older siblings who entered high school during the pre-reform years 1984-87.¹⁶ If we find a positive peer effect, then the pilot scheme can be exploited as an exogenous increase in ε^* as it induced younger siblings to choose Math-Science by having induced their older siblings to opt for Math-Science. Sections 5 and 6 are devoted to analyzing these mechanisms in detail.

The 1988 high school reform can be seen as signaling that combining advanced Math and Science courses is less essential. They do not necessarily need to be bundled together, as post-1988 students can freely combine their advanced courses. It can also be seen as a lowering of the cost of acquiring advanced Math as it does not necessarily have to be combined with advanced Physics or Chemistry. The model thus predicts that this lowering of extrinsic costs (increase in extrinsic benefits) will have a larger effect on the Math-Science choices for younger siblings with Math-Science older siblings.

2.2. Effects of Intrinsic Incentives

Intrinsic incentives, e , are here modeled as primarily varying with older sibling's Math-Science choice. However, intrinsic incentives can also vary by gender or with the strength in social ties. Here

we also assume $\frac{\partial^2 \Delta(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial \varepsilon_j^{*2}} > 0$, which implies that the social multiplier decreases when the cut-off increases. Examining the interaction between changes in extrinsic and intrinsic incentives:

$$\frac{\partial F(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial b \partial e_{old}^j} = \frac{\partial F(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial b \partial \varepsilon_j^*} \frac{\partial \varepsilon_j^*(b, e_{old}^j, \mu)}{\partial e_{old}^j} \quad (7)$$

where the second term is negative, since a higher intrinsic cost means fewer choose Math-Science and the cut-off is lowered. The sign of the first term depends on the change in the density, $\frac{\partial f(\varepsilon_j^*(b, e_{old}^j, \mu))}{\partial \varepsilon_j^*}$, which is positive for values below the unimodal value, $f(0)$, and negative for values above. Thus the first term in (7) is positive (negative) when the cut-off is low (high). This implies that the interaction

¹⁶ Jia and Persson (2014) present a similar model to explain child identity choices of minority couples in China. They primarily exploit exogenous time-series variation in extrinsic incentives induced by the one-child policy, but do not have enough exogenous variation shifting the cut-off in the cross-section and rely on regional variation instead. We rely on time and cross-sectional variation across high schools caused by the policy changes detailed in Section 3.

effect between extrinsic and intrinsic incentives is negative if the share of non-conformers choosing Math-Science is small, $F(\varepsilon_0^*(b, e_{old}^0, \mu))$, but less negative (or even positive) when this share is large. These predictions can be tested by exploiting the fact that the pilot scheme shocked younger siblings' intrinsic incentives indirectly by shocking their older siblings' extrinsic incentives, while the 1988 reform changed their extrinsic incentives directly. We therefore assess the magnitude of this interaction effect by examining younger siblings' Math-Science choices and how they are affected by the social interaction effects with their older siblings.

The model also has strong predictions for gender differences. Math-Science is a more common choice for boys than for girls. If boys have a higher intrinsic cost of not conforming to their older sibling's Math-Science choice or a higher weight put on identity (e.g. because of stereotyping), then the 1988 reform was also a differential shock to intrinsic incentives by gender. Since younger siblings could choose advanced Math without also choosing another advanced Science (i.e. Physics or Chemistry), this may lead to less stereotypical choices.

2.3. Measurement and Empirical Tests

In this section, we empirically test the predictions derived in Sections 2.1 and 2.2. To assess their empirical importance, we quantify how responses to changes in extrinsic incentives vary based on their interaction with (i) the strength of social norms and (ii) intrinsic incentives.

High school curricula can be seen as an extrinsic incentive (or a law) that also sends a signal about school values and the benefits that different course choices are likely to encounter. The pilot scheme rolled out in Danish high schools 1984-87 and the major reform of the curriculum in 1988 can thus be seen as shocks to extrinsic incentives, b , for the exposed cohorts. In this section, we test the model predictions by analyzing the differential impacts of the major 1988 reform.

Intrinsic incentives primarily vary with older sibling's Math-Science choice, e_{old}^j , but their social impact on the Math-Science choice may also vary with social ties. Here we simply measure intrinsic cost differences by the gender of the younger sibling, but in Sections 5 and 6 we further assess heterogeneity by sibling age-difference, gender-composition of sibship, size of sibship, birth-order, family and parental background. In Sections 4-6, we corroborate and exploit how the pilot scheme in 1987-84 exogenously changed younger siblings' intrinsic motivation by exogenously changing older sibling's extrinsic incentives.

The weight placed on identity, μ , varies over peer groups proxied by the Math-Science choices of parents, schoolmates, and schoolmates' older siblings and parents. We obtain exogenous variation in

the number of schoolmates taking advanced Math-Science through variation in how many of their older siblings chose Math-Science and were exposed to the pilot scheme. This way of measuring the social norm should minimize the reflection problem (Manski, 1993, 1995) as older siblings' and parents' relevant educational choices were made years earlier.

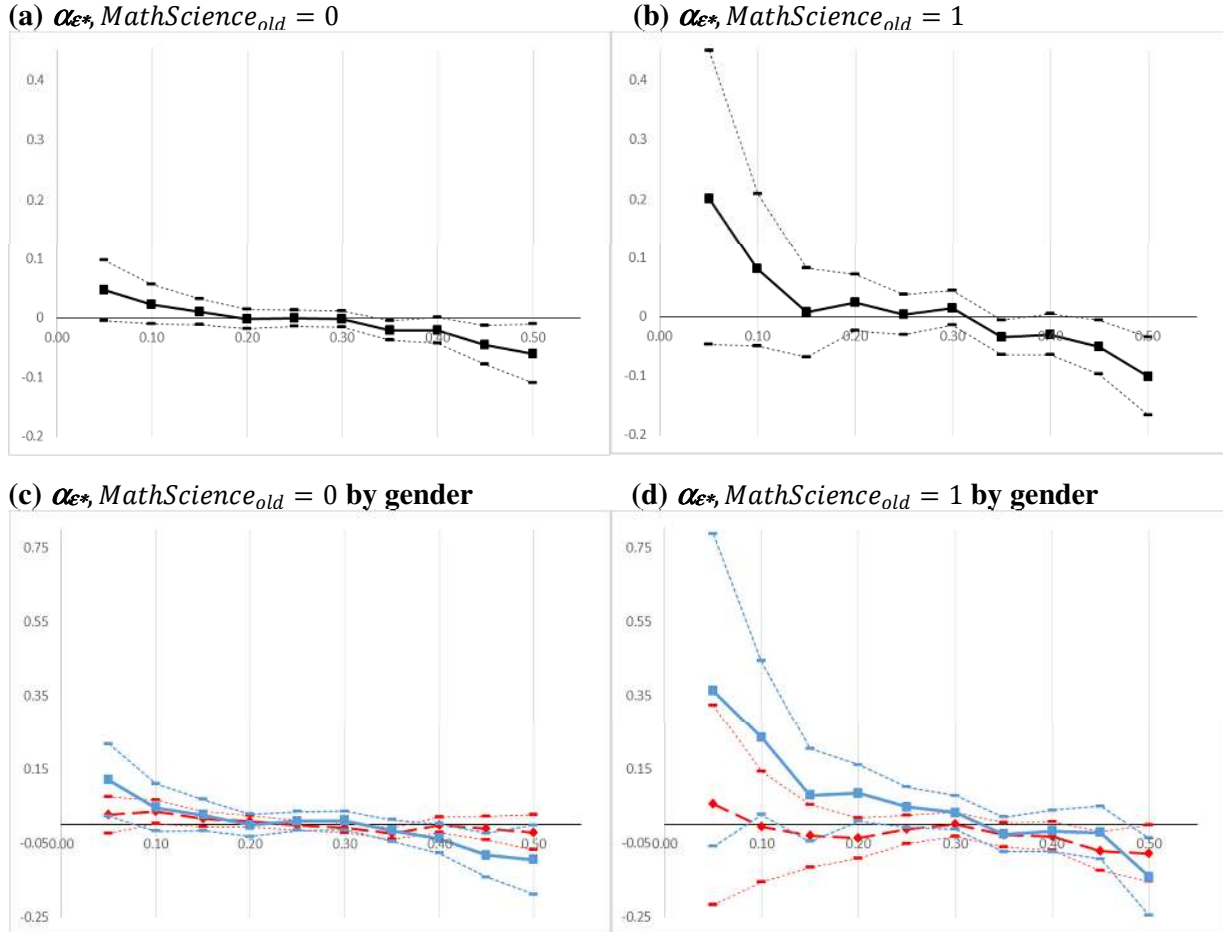
The outcomes and educational choices are thus measured at the individual level, while extrinsic incentives are measured at the school-cohort level (1984-87 pilot) and cohort level (1988 reform), and the social environment is measured both at the sibling, family, and school-cohort level.

To test the prediction in Section 2.1, we need to measure extrinsic incentives and the strength of social norms in the peer group. We measure the fall in extrinsic incentives by the indicator, $I[t \geq 1988]$, which measures how the choice set was universally changed for all high school cohorts entering high school after 1988. The model predicts that the impact of this reform will vary based on the strength of the social norm in the peer group. We cannot observe the cut-off values, ε^* , but proxy them by the fraction of school-cohort-mates' older siblings choosing Math-Science: $\overline{MathScience}_{old,s,t}$, where s denotes high school and t denotes cohort of the younger sibling.¹⁷ Figure 2 presents difference-in-differences coefficient estimates, α_{ε^*} , on the interaction between whether this fraction is higher than a cut-off value, ε^* , ranging from 5% to 50%. That is, the coefficient on the interaction term: $I[\overline{MathScience}_{old,s,t} \geq \varepsilon^*] * I[t \geq 1988]$ in a regression also controlling for whether the younger sibling enters high school after the major reform in 1988, $I[t \geq 1988]$, high school Pilot status, high school fixed effects, and cohort fixed effects plotted against the cut-off, ε^* . Each regression coefficient, α_{ε^*} , thus measures the difference in the impact of the extrinsic incentives post 1988 between high schools above and below the cut-off, ε^* . The reform caused a large decline in the fraction of students choosing Math-Science, and was essentially a decrease in extrinsic incentives as it signals less importance to bundling these advanced courses. That is, a positive regression coefficient, $\alpha_{\varepsilon^*} > 0$, means a negative derivative, $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon} < 0$, of reputational returns. Figure 2 shows how the derivative $\frac{\partial \Delta(\varepsilon)}{\partial \varepsilon}$ goes from being negative when Math-Science is a rare choice in the peer group to being positive when it is a more common choice. Furthermore, a comparison of panels (a) and (b) shows a steeper slope when the older sibling chose Math-Science. This means that the change in the social multiplier is larger and there is more crowding in of the weaker extrinsic incentives for the younger siblings who can conform to their older sibling by choosing Math-Science as they face

¹⁷ We get very similar results if we instead proxy the peer group norm by the fraction of school-cohort-mates' older siblings at a Pilot school or by the fraction of school-cohort-mates' parents with a STEM education.

stronger intrinsic incentives. Panels (c) and (d) show that this is driven by stronger crowding in for younger brothers than for younger sisters. In a model without intrinsic motivation and identity concerns, we would expect identical individuals to respond identically to universal curricula change independently of their peer group.

Figure 2. Crowding in of Extrinsic Incentives.



Note: The figure displays the regression coefficient, α_{ϵ^*} , and 95% confidence intervals (vertical axis) by the fraction of school-cohort-mates' older siblings with Math-Science, $\overline{MathScience}_{old,s,t}$ (horizontal axis) as a proxy for the variation in the cut-off value, ϵ^* , for whether younger siblings are indifferent between choosing Math-Science or not. All estimates are displayed separately by older siblings' Math-Science choice and by gender in panels (c) and (d), where the light solid line with squares denotes boys and the dark dashed line with diamonds denotes girls.

To test the prediction in Section 2.2 we also need to measure the intrinsic cost. As argued previously, we hypothesize that the intrinsic cost of not conforming is higher for boys, since gender is a salient feature of identity and boys may face stronger social pressure to conform as math and science are stereotypical male choices. We find that the drop in Math-Science skill supply after the 1988 reform is 8.7 percentage points larger for boys than for girls.

Overall, this section has shown how the interplay between changes in extrinsic incentives and their interactions with intrinsic costs and social norms across peer groups have substantially contributed to the continuous fall in the Math-Science skill supply. Over the period 1980-2000, this advanced skill supply has decreased by 50% among boys and the gender gap has closed by 57% as the social pressure to take advanced math and science in high school tends to have been particularly strong for boys.

In order to better understand why we observe more crowding in for boys and younger siblings facing a stronger STEM norm, we turn to better understanding the underlying mechanisms by estimating causal sibling spillover effects. These are essential for understanding how strongly social multipliers may amplify or dampen impacts of education policies depending on the composition of the peer group.

3. Identification of Peer Effects Using a High School Pilot Scheme

We exploit some unique features and changes in institutions in Denmark to identify the peer effects in sibling interactions. 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 peer 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.

3.1. Identifying Peer 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.¹⁸ These identification issues can be illustrated in a model which is

¹⁸ Manski (1993; 1995) provides a more complete and general analysis of the identification of peer effects (or more generally endogenous effects), while Moffitt (2001) introduces the conceptual framework we adopt here. In a more recent contribution, Angrist (2014) discusses the identification challenges.

linear in the peer effect. We assume, without loss of generality, that there are only two individuals in each peer group - an older sibling and a younger sibling.¹⁹

$$MathScience_{old} = \pi_0 + \pi_1 MathScience_{young} + \pi_2 X_{old} + \pi_3 X_{young} + \pi_4 a_f + \varepsilon_{old,f} \quad (8)$$

$$MathScience_{young} = \beta_0 + \beta_1 MathScience_{old} + \beta_2 X_{old} + \beta_3 X_{young} + \beta_4 a_f + \varepsilon_{young,f} \quad (9)$$

where $MathScience_i$ denotes whether sibling i chose advanced math with an advanced science course in high school, X_i denotes observable characteristics of sibling i , a_f denotes sibling pair specific characteristics like family background, gender composition, and age difference. Finally, $\varepsilon_{i,f}$ denotes other unobserved factors affecting the $MathScience$ choice of individual i in sibling 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 (9) 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 making similar high school course choices. Omitted variables bias due to contextual effects arises if we are not able to observe all these relevant sibling pair specific (a_f) and individual variables (X_i). 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.²⁰ Nevertheless, this is not a necessary exclusion restriction as our empirical strategy addresses all these three empirical concerns, since the exogenous variation in the cost of acquiring advanced math and science for the older sibling is independent of both sibling pair specific factors and individual sibling characteristics.

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

¹⁹ It is straightforward to generalize this setting to larger peer groups. Brock and Durlauf (2001) discuss identification in nonlinear peer effects models.

²⁰ The developmental psychology literature supports that the direction of behavioral influence goes from the older sibling to the younger sibling (Buhrmester, 1992). Altonji et. al (2013) also corroborate this assumption and impose it as an identifying assumption to estimate causal sibling influences on adolescence substance use.

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 was introduced unexpectedly just before the older sibling made the choice of advanced high school courses. The pilot scheme thus provides us with exogenous variation in the cost of acquiring advanced math and science for the older sibling (captured by $PilotIntro_{old}$) that does not directly influence the younger sibling *and* is independent of any sibling pair specific (a_f) and individual variables (X_i). Substituting this into (8) and (9) we get:

$$MathScience_{old} = \pi_0 + \gamma PilotIntro_{old} + \pi_2 X_{old} + \pi_3 X_{young} + \pi_4 a_f + \varepsilon_{old,f} \quad (10)$$

$$MathScience_{young} = \beta_0 + \beta_1 MathScience_{old} + \beta_2 X_{old} + \beta_3 X_{young} + \beta_4 a_f + \varepsilon_{young,f} \quad (11)$$

Younger siblings attend high school in a regime, where they have an even more flexible curriculum as advanced math and science courses can be combined as they like - the main requirement is that they choose at least two (and at most three) optional advanced courses. This particular institutional setting thus provides us with a unique quasi-experiment for identifying peer effects in math and science - going from the older sibling's course choice to the younger sibling's course choice. We can thus interpret the IV estimate of β_1 in the structural equation (11) as capturing this causal peer effect when the first-stage equation (10) includes $PilotIntro_{old}$ as an instrument for $MathScience_{old}$ which endogenously affects $MathScience_{young}$, because the instrument only affects the older sibling directly and the younger one merely through endogenous social interaction.²¹ The identifying assumptions are corroborated in Joensen and Nielsen (2009, 2014) 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 second high school year. Furthermore, the instrument has a strong influence on the choice of math and science courses for the older sibling. We return to these empirical issues in Section 5. 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.

²¹ Moffitt (2001) labels this type of identification strategy as a partial-population policy intervention.

3.2. *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.²² 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.

This regime implied that students upon high school graduation would have achieved one of three math levels available: advanced, intermediate, or basic level. The difference between the three levels is reflected in the number of lessons per week, as well as in the content of the courses. For instance, the extent of geometry and algebra increases as the level becomes more advanced. In the empirical analysis, we focus on whether students choose advanced math and science, meaning that the intermediate and basic level courses are lumped together. 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 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.

3.3. *The Pilot Scheme*

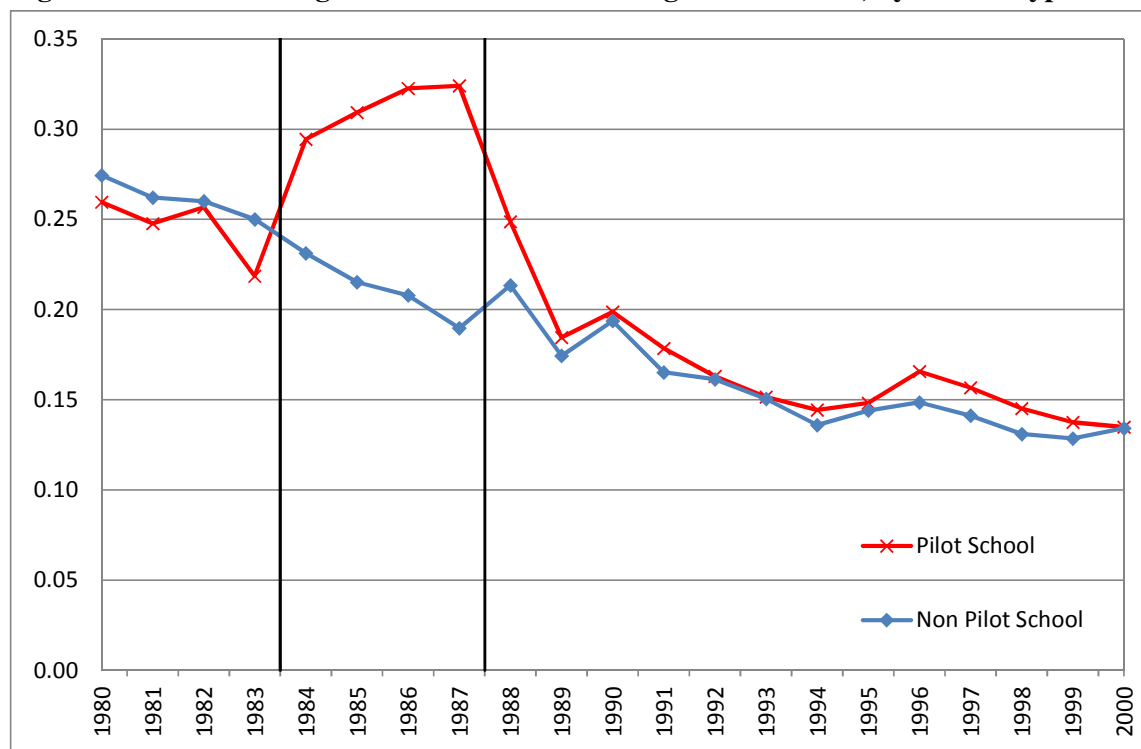
The pilot scheme was implemented as an experimental curriculum at about half of the high schools prior to the 1988-reform. The purpose of the pilot scheme was to test the impact of increased flexibility prior to the 1988-reform. Figure 3 illustrates the consequences of the pilot scheme on the course packages of the high school youth. Prior to the pilot scheme, the fraction choosing advanced math and science declined and went 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.

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 ($PilotSchool=0, PilotIntro=1$),

²² 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.

and schools where the pilot scheme was implemented prior to enrollment of the relevant cohort ($PilotSchool=1, PilotIntro=0$).

Figure 3. Fraction of High School Cohorts Choosing Math-Science, by School Type



Note: Pilot schools include all schools with pilot status at any point in time during 1984-87; i.e. 64 schools in total.

Schools were not randomly assigned to become pilot schools. Instead, from 1984-86, they could apply to the Ministry of Education for permission to adopt the experimental curriculum, whereas in 1987 the high school principals could make this decision without approval from the ministry.²³ It is not possible to directly test whether the pilot schools represent a sample of schools which is essentially random with respect to math ability, but we corroborate that this is a reasonable assumption.

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 high school ($PilotSchool=0, PilotIntro=1$) and those who knew that the school was a pilot school before they applied for entering the school ($PilotSchool=1, PilotIntro=0$).

²³ 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.

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	Schools	N	MathScience Schools	Schools	N	MathScience Schools	Schools
1984	12,758	0.2482	122	3,145	0.3253	22	0	0.0000	0	15,903	0.2634	144
1985	10,645	0.2330	107	1,802	0.3330	15	3,069	0.3271	22	15,516	0.2632	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	0.2243		8,201	0.3159		15,591	0.3360		64,489	0.2630	

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 who are traditional high schools only offering advanced Math with advanced Physics (PilotSchool=0), who are unexpectedly introducing the pilot scheme combining advanced Math with advanced Chemistry (PilotSchool=0, PilotIntro=1) and who already have adopted the pilot scheme (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.²⁴ 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 participating schools is exogenous with respect to student ability, 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 then introduces the experimental curriculum for the first time, and it takes the value zero otherwise. This instrument is valid 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. This assumption is violated if the school decides to participate in the program based on the math abilities of local students. In Section 4 below, we test for similarities of student and parent characteristics across school status, and we find almost no significant differences in characteristics determined prior to high school. The assumption is also violated 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. Such effects would influence the younger siblings if they attend the same high school as their older siblings, and the effects could confound the sibling peer effects.

²⁴ 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.

However, Figures 1 and 3 showed that the pilot scheme was introduced following a declining trend in the fraction choosing advanced math and science, and therefore qualified teachers would most likely be available. Furthermore, the relevant compliers would switch from intermediate chemistry and intermediate math to advanced chemistry and advanced math which would require only 5-6 additional weekly lectures for one class per cohort per school.

In the empirical analysis we investigate to what extent the spillover effects may be going through the high school rather than through sibling interactions. We find that the bulk (at least 60%) of the spillover effect goes through social interactions among siblings.

Table 2 presents placebo tests to support instrument exogeneity. 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 (1.5 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 3. 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.

The instrument is strong if the unexpected introduction of the pilot scheme induces students to choose advanced math and science, which is directly tested and validated in Section 4. The instrument satisfies the monotonicity (or uniformity) condition if individuals who chose advanced math and science when required to do advanced physics and intermediate chemistry would also have chosen advanced math and science if they unexpectedly got the option of replacing advanced physics with advanced chemistry and replacing intermediate chemistry with intermediate physics. 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.

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.

Table 2. Placebo Tests on Pre-Pilot Cohorts

	Parameter Estimates (Standard Errors)		
	(1)	(2)	(3)
	First-stage: <i>MathScience</i>		
<i>Placebo PilotIntro</i>	-0.015 ** (0.006)	0.002 (0.006)	-0.005 (0.005)
Pilot Intro cohort (actual)	1984-1985	1984-1986	1984-1987
Pilot Intro cohort (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.

3.4. 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.²⁵ In particular, the mathematical 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 educational 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.

²⁵ 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.²⁶ 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 the 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.²⁷ Thus $MathScience_{young}$ in equation (8) is an indicator for whether the younger sibling chooses to combine advanced math with either advanced physics or advanced chemistry.

4. Data Description

4.1. Sample Selection

For our empirical analysis we use a panel data set comprising the population of individuals starting high school from 1984 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

²⁶ 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.

²⁷ 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 about 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. 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.²⁸ In our main analysis, we focus on a homogeneous sample of relatively closely spaced sibling pairs (cohorts 1988-91, age gap ≤ 4 years).²⁹

4.2. 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 21.6% 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.

We also include entry cohort fixed effects, sibling gender composition, parental background, and high school specific controls. Parental background includes 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. We leave out post-graduation control variables and thus estimate the total effect of advanced math and science.

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

²⁸ 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.

²⁹ 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.

had already adopted the pilot program ($PilotSchool=1$, $PilotIntro=0$) previous 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 validity of $PilotIntro$ as an instrument for older siblings' course package choice.

5. Estimates of Sibling Peer Effects

Table 3 presents the main results from the empirical analysis of how an increase in extrinsic incentives for the older sibling spill over on their younger sibling through increasing their intrinsic motivation. We present the results from OLS and 2SLS while the main results are similar for alternative methods.³⁰ The OLS regressions indicate a strong positive association between math and science course choices of older and younger siblings. The reduced form estimates suggest that there are spillover effects from the introduction of the pilot scheme to the younger siblings, while the IV estimates suggest that the effect goes through the older siblings' course choices. However, the effects are only statistically significant when the age distance between the siblings is less than four years.³¹ When the age distance is less than four years, the magnitude of the estimates ranges from 0.35-0.52, which suggests a very strong peer effect.

Including additional control variables does not significantly affect the point estimates, lending additional support to the exclusion restriction and exogeneity of $PilotIntro$ which is imperative for the causal inference based on the IV estimates. Columns (3) and (6) add explanatory variables related to high schools: an indicator variable for whether the siblings attended the same high school and predetermined high school means of parents' highest completed education and income. These mean variables are thought to approximate permanent high school specific effects such as the quality of science teachers or the expertise in science teaching. In Table B8 in Appendix B we further investigate

³⁰ We report the results from employing the bivariate probit estimator in Table B6 in Appendix B. Our main results are broadly robust and the first-stages are literally unchanged. Conclusions from the bivariate probit model appear slightly stronger; particularly in the subgroup analysis. However, an assessment along the lines of Altonji, Elder and Taber (2005) reveals that primarily the parametric assumptions, and not the exogenous variation, drive identification when covariates are included in the bivariate probit model. Therefore, we have chosen to report the results from linear probability models as our main results. However, the disadvantage from using linear probability models is that a fraction of the predicted outcome probabilities are negative (around 20%). In addition, we have used the semi-parametric estimator by Abadie (2003) as a robustness check. The second stage estimate without control variables is almost identical to the bivariate probit (marginal effect 0.384 vs. 0.360).

³¹ Table B7 in Appendix B presents the results when the maximum age difference is held constant and the cohorts are allowed to vary. These results confirm that the results vary with age difference and not with cohorts. This is consistent with sibling pairs with an age difference of five years or more being considered separate sibships (Adams, 1972).

Table 3. Estimates of Peer Effects: Main Results

		Parameter Estimates (Standard Errors)					
		OLS			2SLS		
	N	(1)	(2)	(3)	(4)	(5)	(6)
First-stage: Older Sibling MathScience							
<i>PilotIntro</i>					0.068 ***	0.071 ***	0.065 ***
Younger Sibling 1988-91, ≤4y	7,786				(0.017)	(0.016)	(0.017)
<i>PilotIntro</i>					0.065 ***	0.070 ***	0.067 ***
Younger Sibling 1988-92, ≤5y	10,571				(0.015)	(0.014)	(0.014)
<i>PilotIntro</i>					0.060 ***	0.067 ***	0.063 ***
Younger Sibling 1988-93, ≤6y	12,717				(0.013)	(0.013)	(0.013)
<i>PilotIntro</i>					0.067 ***	0.073 ***	0.069 ***
Younger Sibling 1988-97, ≤10y	17,691				(0.011)	(0.010)	(0.011)
Reduced-form: Younger Sibling MathScience							
<i>PilotIntro</i>		0.036 **	0.032 **	0.023			
Younger Sibling 1988-91, ≤4y		(0.015)	(0.014)	(0.014)			
<i>PilotIntro</i>		0.020 *	0.017	0.01			
Younger Sibling 1988-92, ≤5y		(0.012)	(0.012)	(0.012)			
<i>PilotIntro</i>		0.017	0.013	0.067			
Younger Sibling 1988-93, ≤6y		(0.011)	(0.011)	(0.011)			
<i>PilotIntro</i>		0.010	0.009	0.006			
Younger Sibling 1988-97, ≤10y		(0.009)	(0.009)	(0.009)			
Outcome: Younger Sibling MathScience							
Older Sibling MathScience		0.140 ***	0.147 ***	0.145 ***	0.523 **	0.447 **	0.346
Younger Sibling 1988-91, ≤4y		(0.009)	(0.010)	(0.010)	(0.231)	(0.205)	(0.221)
Older Sibling MathScience		0.145 ***	0.151 ***	0.150 ***	0.314	0.241	0.149
Younger Sibling 1988-92, ≤5y		(0.008)	(0.008)	(0.008)	(0.192)	(0.168)	(0.177)
Older Sibling MathScience		0.142 ***	0.149 ***	0.147 ***	0.276	0.193	0.106
Younger Sibling 1988-93, ≤6y		(0.007)	(0.007)	(0.007)	(0.183)	(0.155)	(0.168)
Older Sibling MathScience		0.139 ***	0.146 ***	0.144 ***	0.143	0.121	0.084
Younger Sibling 1988-97, ≤10y		(0.006)	(0.006)	(0.006)	(0.133)	(0.118)	(0.127)
Control Variables:							
Entry Cohort Fixed Effects			+	+		+	+
Sibling pair gender composition			+	+		+	+
Sibling pair in same High School				+			+
<i>Parental variables (for mother and father):</i>							
Age, Highest Completed Education, and Income			+	+		+	+
HS Mean of Highest Completed Education and Income				+			+

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

whether the spillover works through the school rather than the sibling. We present results from a placebo test where we randomly match only children from entry cohorts 1984-87 with only children

from entry cohorts 1988-91. At the left hand side of the table we match the children randomly and at the right hand side we randomly match only children attending the *same* high school. No matter how we match the children, there is no correlation between their course choices. When we match only children attending the same high schools, the reduced-form coefficient and the 2SLS coefficients are larger than when they are randomly matched. However, the estimates are 40% smaller compared to the main results in Table 3 and insignificant. When high school specific variables are added (columns (6) and (8)) they are even smaller. We conclude that it is important to account for the high school specific variables in order to rule out that the spillover effect goes through the school.³²

The IV point estimates are larger than the OLS estimates, although not significantly so. However, it suggests that 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 an always taker (“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.

The fact that the peer effect is stronger when the age difference is limited to four years than otherwise, could suggest that a smaller age difference implies closer ties. However, it could also reflect more sibling rivalry or stronger role model effects among closely spaced siblings. Unfortunately, the data does not allow us to draw more detailed inference about the importance of sibling spacing closer than 4 years, but we examine heterogeneity in detail to inform potential mechanisms in the next section.

6. Understanding Heterogeneity in Sibling Interactions

In this section, we seek to better understand heterogeneity in peer effects. We examine which older siblings respond more (or less) to changes in extrinsic incentives, as well as when the shock to the younger sibling’s intrinsic motivation matters more (or less). That is, for which sibling pairs the peer effect is stronger (or weaker). We explore differences in complier characteristics and peer effects across gender, sibship composition, ability, parental background, and strength of ties. These heterogeneous effects lead towards inference about plausible causal mechanisms.

Our theoretical framework in Section 3 showed how educational choices are shaped by peers and the social environment, but was general enough to encompass a variety of underlying mechanisms. The student may gain utility from behaving similarly to an older sibling who inspires academic

³² 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. When we leave out the placebo siblings within the same high school with an age difference of less than two years, then the unconditional reduced form coefficient in (5) further falls to insignificant 0.014. Likewise, the 2SLS estimate is also insignificant and 60% smaller compared to our main estimates in Table 3.

aspirations, and the student may compete with the older sibling in terms of academic achievements in order to improve his or her reputational return in the peer group. 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. Another possible influence of the environment may be productivity spillover effects in which case a student with an older sibling who did math and science may be able to perform better in school due to assistance with homework and learning. Finally, there may be externalities due to joint leisure in which case a student with a peer sharing his or her passion for math and science would be able to share public goods with this peer (e.g. nerdy jokes and movies) or to perceive the academic culture as relatively more welcoming.

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 (Black et al., 2005; Björklund and Jäntti, 2012). Adams (1972) suggests that second and middle children would often try to catch up with the first child and thus compete, but the youngest child less so. Conley (2000) stressed that same-sex sibling 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.³³ The math discipline attracts 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

³³ See e.g. Altonji (1995), Joensen and Nielsen (2009, 2014), Cortes et al. (2014), and Falch et al. (2014).

math and science high school track is explained by gender differences in competitiveness. Therefore, *MathScience* course choice may be a battle field 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.

6.1. 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 is around 1.5 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 are around 50% more likely to comply to the unexpected introduction of the pilot scheme. 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 5-6 percentage points. The sibling spillover effects also differ by gender: the point estimates are larger for older brothers and for younger brothers. This may support a hypothesis of sibling rivalry typically found to be systematically stronger among pairs of brothers.³⁴ 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 6.3.

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 at the end of high school. We caution that this measure may be affected by the course choice but it is the best we can do with the available data. 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 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

³⁴ When we divide the sample into four subgroups (brother-brother, brother-sister, sister-brother, sister-sister), we find that the causal peer effect is largest and only significantly positive for brother-brother pairs.

Table 4. Estimates of Peer Effects: Heterogeneity by Gender

	Parameter Estimates (Standard Errors)											
	Older Sibling						Younger Sibling					
	Brother			Sister			Brother			Sister		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	First-stage: Older Sibling MathScience											
<i>PilotIntro</i>	0.105 *** (0.027)	0.103 *** (0.027)	0.095 *** (0.028)	0.046 ** (0.020)	0.045 ** (0.020)	0.041 ** (0.020)	0.080 *** (0.026)	0.087 *** (0.024)	0.077 *** (0.025)	0.058 ** (0.023)	0.056 ** (0.022)	0.051 ** (0.023)
Relative to overall first-stage	1.544	1.451	1.462	0.676	0.634	0.631	1.176	1.225	1.185	0.853	0.789	0.785
	Reduced-form: Younger Sibling MathScience											
<i>PilotIntro</i>	0.047 ** (0.022)	0.046 ** (0.021)	0.033 (0.021)	0.026 (0.020)	0.020 (0.019)	0.013 (0.019)	0.062 ** (0.025)	0.062 ** (0.025)	0.047 * (0.026)	0.007 (0.015)	0.005 (0.015)	0.001 (0.015)
	Outcome: Younger Sibling MathScience											
Older Sibling <i>MathScience</i>	0.452 ** (0.216)	0.442 ** (0.212)	0.352 (0.226)	0.577 (0.462)	0.440 (0.436)	0.310 (0.471)	0.777 ** (0.366)	0.716 ** (0.324)	0.607 * (0.354)	0.123 (0.250)	0.092 (0.256)	0.021 (0.291)
Control Variables:												
Entry Cohort Fixed Effects		+	+		+	+		+	+		+	+
Sibling pair gender composition		+	+		+	+		+	+		+	+
Sibling pair in same High School			+			+			+			+
Parental variables (for mother and father):												
Age, Highest Completed Education, and Income		+	+		+	+		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+			+			+
Mean of Older Sibling <i>MathScience</i>		0.483			0.189			0.343			0.316	
Mean of Younger Sibling <i>MathScience</i>		0.194			0.186			0.310			0.091	
Number of Sibling Pairs		3,673			4,113			3,490			4,296	

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table 5. Estimates of Peer Effects: Heterogeneity by Ability

	Parameter Estimates (Standard Errors)					
	Older Sibling					
	GPA < P50		P50 < GPA < P90		GPA > P90	
	(1)	(2)	(3)	(4)	(5)	(6)
	First-stage: Older Sibling <i>MathScience</i>					
<i>PilotIntro</i>	0.086 *** (0.023)	0.068 *** (0.023)	0.040 (0.028)	0.050 * (0.027)	0.135 ** (0.061)	0.149 ** (0.060)
Relative to overall first-stage	1.265	1.046	0.588	0.769	1.985	2.292
	Reduced-form: Younger Sibling <i>MathScience</i>					
<i>PilotIntro</i>	0.035 * (0.019)	0.020 (0.019)	0.059 ** (0.024)	0.039 * (0.024)	-0.051 (0.051)	-0.027 (0.051)
	Outcome: Younger Sibling <i>MathScience</i>					
Older Sibling <i>MathScience</i>	0.408 * (0.238)	0.291 (0.281)	1.452 (1.069)	0.785 (0.562)	-0.376 (0.440)	-0.181 (0.357)
Control Variables:						
Entry Cohort Fixed Effects		+		+		+
Sibling pair gender composition		+		+		+
Sibling pair in same High School		+		+		+
Parental variables (for mother and father):						
Age, Highest Completed Education, and Income		+		+		+
HS Mean of Highest Completed Education and Income		+		+		+
Mean of Older Sibling <i>MathScience</i>		0.290		0.347		0.420
Mean of Younger Sibling <i>MathScience</i>		0.170		0.208		0.210
Number of Sibling Pairs		3,768		3,142		876

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively

(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.³⁵

6.2. Birth Order and Size of Sibship

In Table 6 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 he or she 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.

³⁵ 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 imprecise.

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. Appendix C provides 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 effect 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). We find a strong positive direct effect of the first- on the second-born, but negative direct effects on the third-born. Our estimates of the direct effects are too noisy to draw strong inference, but 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.

6.3. *Strength of Ties between Siblings*

The strength of ties relate to the nature and the duration of the relationship as well as the frequency and intensity of interactions. We explore heterogeneity in peer effects depending on whether the sibling pairs share both parents (98% of our sample) and whether the sibling pairs grew up together (96% of our sample). However, because these groups comprise most of the sample, the results are almost identical to the main results. Finally, we also examine heterogeneity by whether the sibling pair attended the same high school or not. Peer effects are only significant for siblings that attend different high schools. These results are shown in Table B9 in Appendix B.

6.4. *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 compatible educational 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 at high school entrance 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

Table 6. Estimates of Peer 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)
	First-stage: Older Sibling MathScience											
<i>PilotIntro</i>	0.147 *** (0.037)	0.143 *** (0.037)	0.048 ** (0.019)	0.043 ** (0.019)	0.032 (0.031)	0.035 (0.030)	0.085 *** (0.021)	0.078 *** (0.020)	0.069 *** (0.025)	0.052 ** (0.024)	0.066 *** (0.024)	0.075 *** (0.023)
Relative to overall first-stage	2.162	2.200	0.706	0.662	0.471	0.538	1.250	1.200	1.015	0.800	0.971	1.154
	Reduced-form: Younger Sibling MathScience											
<i>PilotIntro</i>	0.016 (0.032)	0.007 (0.031)	0.041 ** (0.016)	0.026 (0.016)	0.031 (0.027)	0.006 (0.026)	0.038 ** (0.017)	0.029 * (0.017)	0.033 (0.021)	0.021 (0.020)	0.038 * (0.021)	0.020 (0.020)
	Outcome: Younger Sibling MathScience											
Older Sibling <i>MathScience</i>	0.107 (0.213)	0.051 (0.213)	0.839 * (0.434)	0.613 (0.423)	0.988 (1.174)	0.160 (0.725)	0.449 ** (0.216)	0.375 * (0.222)	0.473 (0.316)	0.411 (0.398)	0.577 * (0.347)	0.265 (0.264)
Control Variables:												
Entry Cohort Fixed Effects		+		+		+		+		+		+
Sibling pair gender composition		+		+		+		+		+		+
Sibling pair in same High School		+		+		+		+		+		+
Parental variables (for mother and father):												
Age, Highest Completed Education, and Income		+		+		+		+		+		+
HS Mean of Highest Completed Education and Income		+		+		+		+		+		+
Mean of Older Sibling <i>MathScience</i>	0.295		0.337		0.325		0.329		0.343		0.314	
Mean of Younger Sibling <i>MathScience</i>	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: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table 7. Estimates of Peer Effects: Heterogeneity by Parental Education

	Parameter Estimates (Standard Errors)											
	Father						Mother					
	STEM (narrow)		STEM (broad)		> 4-year college		STEM (narrow)		STEM (broad)		> 4-year college	
	0	1	0	1	0	1	0	1	0	1	0	1
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
	First-stage: Older Sibling MathScience											
<i>PilotIntro</i>	0.068 *** (0.019)	0.048 (0.036)	0.067 *** (0.019)	0.056 (0.035)	0.046 ** (0.019)	0.136 *** (0.038)	0.059 *** (0.017)	0.184 ** (0.091)	0.062 *** (0.017)	0.097 (0.084)	0.061 *** (0.017)	0.140 * (0.072)
Relative to overall first-stage	1.046	0.738	1.031	0.862	0.708	2.092	0.908	2.831	0.954	1.492	0.938	2.154
	Reduced-form: Younger Sibling MathScience											
<i>PilotIntro</i>	0.009 (0.016)	0.056 * (0.033)	0.015 (0.016)	0.038 (0.031)	0.012 (0.016)	0.066 ** (0.032)	0.019 (0.014)	0.071 (0.083)	0.016 (0.014)	0.160 ** (0.074)	0.013 (0.015)	0.181 ** (0.062)
	Outcome: Younger Sibling MathScience											
Older Sibling <i>MathScience</i>	0.139 (0.225)	1.166 (1.030)	0.219 (0.234)	0.689 (0.654)	0.253 (0.344)	0.490 * (0.251)	0.329 (0.248)	0.382 (0.429)	0.257 (0.231)	1.648 (1.381)	0.214 (0.237)	1.289 * (0.706)
Control Variables:												
Entry Cohort Fixed Effects	+	+	+	+	+	+	+	+	+	+	+	+
Sibling pair gender composition	+	+	+	+	+	+	+	+	+	+	+	+
Sibling pair in same High School	+	+	+	+	+	+	+	+	+	+	+	+
Parental variables (for mother and father):												
Age, Highest Completed Education, and Income	+	+	+	+	+	+	+	+	+	+	+	+
HS Mean of Highest Completed Education and Income	+	+	+	+	+	+	+	+	+	+	+	+
Mean of Older Sibling <i>MathScience</i>	0.298	0.428	0.302	0.406	0.322	0.347	0.323	0.425	0.323	0.398	0.326	0.352
Mean of Younger Sibling <i>MathScience</i>	0.168	0.262	0.172	0.242	0.183	0.210	0.185	0.283	0.186	0.245	0.188	0.213
Number of Sibling Pairs	5,984	1,802	5,828	1,958	5,958	1,828	7,412	374	7,309	477	7,232	554

Note: The narrow definition of STEM fields follows the definition by the US Department of Homeland Security (DHS), which almost entirely includes Math, Engineering, Natural and Technical Sciences, while the broad definition follows the definition by the National Science Foundation (NSF), which also includes some social sciences and life sciences. Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table 8. Estimates of Peer 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)
	First-stage: Older Sibling MathScience							
<i>PilotIntro</i>	0.037 *	0.031	0.060	0.065	0.114 ***	0.103 ***	0.088	0.065
	(0.021)	(0.022)	(0.044)	(0.045)	(0.031)	(0.032)	(0.054)	(0.055)
Relative to overall first-stage	0.804	0.756	1.304	1.585	1.086	1.084	0.838	0.684
	Reduced-form: Younger Sibling MathScience							
<i>PilotIntro</i>	0.027	0.008	0.009	0.014	0.026	0.014	0.114 **	0.089 *
	(0.021)	(0.021)	(0.045)	(0.044)	(0.024)	(0.024)	(0.047)	(0.046)
	Outcome: Younger Sibling MathScience							
Older Sibling <i>MathScience</i>	0.739	0.253	0.144	0.222	0.227	0.133	1.296	1.377
	(0.671)	(0.674)	(0.743)	(0.666)	(0.209)	(0.223)	(0.864)	(1.241)
Control Variables:								
Entry Cohort Fixed Effects		+		+		+		+
Sibling pair gender composition		+		+		+		+
Sibling pair in same High School		+		+		+		+
Parental variables (for mother and father):								
Age, Highest Completed Education, and Income		+		+		+		+
HS Mean of Highest Completed Education and Income		+		+		+		+
Mean of Older Sibling <i>MathScience</i>		0.164		0.268		0.451		0.567
Mean of Younger Sibling <i>MathScience</i>		0.159		0.274		0.174		0.245
Number of Sibling Pairs		3,137		976		2,661		1,012

Note: This table applies the narrow definition of STEM fields which follows the definition by the US Department of Homeland Security (DHS). Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

to helping the students deciding what to do after compulsory school.³⁶ A recent survey showed that 79% of students in 9-10th grade answered that their parents were closely involved in their high school enrollment choice. We have no reason to believe that parental involvement is lower in our sample. This suggests that students discuss their educational 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 also 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 will take 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 7, we investigate heterogeneity of the incentive effects by parents' education. We define indicators for whether the mother or the father, respectively, has completed at least a 4-year college degree. Similarly, we construct indicator variables for whether the mother or the father has completed a STEM field education. We use a narrow definition of STEM fields which follows the definition by the US Department of Homeland Security (DHS), which almost exclusively includes Math, Engineering, Natural and Technical Sciences, while the broad definition follows the definition by the National Science Foundation (NSF), which also includes some social sciences and life sciences.

We find that the older sibling tends to respond more to the pilot scheme when either of the parents has completed a college education or when the mother has completed a STEM field education. If the father has completed a STEM field education, the older sibling does not significantly react to the introductions of the pilot scheme. This is most likely because the potential is exhausted as 42.8% (narrow definition) and 40.6% (broad definition) of the high school students in such families already choose math and science (see the bottom panel of Table 7). When we divide by gender of the older sibling in Table 8, it is evident that older daughters respond more strongly to extrinsic incentives in

³⁶ See Ministry of Education (2013).

families where at least one of the parents has a STEM field education. This is again because there is a large pool of unexploited talent because only 27% of the girls in such families choose math and science (see the bottom panel of Table 8) and the marginal monetary payoff for them is thus expected to be high.³⁷ On the other hand, older sons are less likely to be compliers if at least one of their parents are in a STEM field, but this is again likely to be because these sons are very likely to be always-takers (57% of them choose math and science) and do accordingly not face a high expected marginal monetary payoff (Joensen and Nielsen, 2014). These results are even stronger if we single out fathers and older sons (not tabulated). 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 intrinsic sibling spillover.

The estimates of the spillover effects are rarely significant in the subsamples, but the point estimates are systematically higher when the parents have a college degree or a STEM field education. It is not evident that this necessarily reflects that peer effects are higher in these instances. It may reflect that the parents are deeply 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 contents 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 path-breakers by providing information on course content.

7. Conclusion

We find evidence that social interactions among siblings are important for skill formation. Older siblings' educational choices causally influence younger siblings' choices. Younger siblings are 3.5% 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. This implies a sizeable peer effect of almost 50 percentage points which varies across subgroups. We argue that the pattern of the results across gender, ability, birth spacing and birth order implies that the most likely mechanism is competition and sibling rivalry.

We provide a framework for understanding the mechanisms behind the large peer effects among siblings which focuses on the interplay between extrinsic and intrinsic incentives. We show that the combination of a major reform where course choices are made extremely flexible (extrinsic incentives) and positive peer effects (intrinsic incentives and identity concerns), may lead to a continuously

³⁷ This is analyzed in detail in Joensen and Nielsen (2014).

declining supply of advanced math and science skills in social environments with a strong math-science norm and high intrinsic costs of not conforming. We argue that this is what caused the fall in math-science skill supply and the gender convergence in Denmark since the major high school reform in 1988. This exemplifies how skill supply is shaped by institutions and their interactions with the social environment.

Our results imply that social interactions in the family exacerbate the inequality across households. First, the presence of possibly large sibling peer effects in educational choices reveals that the strong sibling correlations conceal influential interactions among siblings. Second, we find systematically larger sibling spillover effects in families where either of the parents has a college education or a STEM field education. This suggests that parents are closely involved in the educational decision and a part of the social environment in which these decisions are made. However, more hard evidence is needed about the role of the family and siblings in human development more generally.

References

- Abadie, A. (2003), Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics* 113: 231 – 263.
- Adermon, A. (2013), *Essays on the Transmission of Human Capital and the Impact of Technological Change*. Department of Economics, Uppsala University. Economics Studies 135. ISBN 978-91-85519-42-2. ISSN 0283-7668.
- 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 (2013), Identifying Sibling Influence on Teenage Substance Use. IFS WP 13/04.
- 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.
- Angrist, J. (2014), The perils of peer effects. *Labour Economics*.30: 98-108.
- 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.
- Babcock, L., and Laschever, S. (2009). *Women don't ask: Negotiation and the gender divide*. Princeton University Press.
- 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.

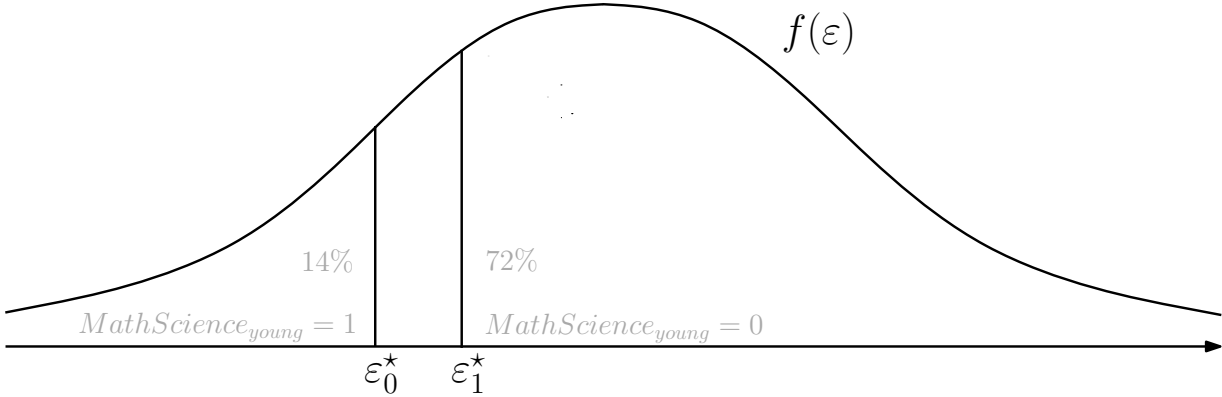
- Becker, G. S. and N. Tomes (1986), Human capital and the rise and fall of families. *Journal of Labor Economics* 4(3): S1-S39.
- Benabou, R. and J. Tirole (2006), Incentives and Prosocial Behavior. *American Economic Review* 96(5): 1652-1678.
- Benabou, R. and J. Tirole (2011), Laws and Norms. *NBER WP #17579*.
- Bertrand, M. (2010), New perspectives on gender. In: Ashenfelter, O. and D. Card (Eds.), *Handbook of Labor Economics*, North-Holland: Amsterdam, 1545-1592.
- Bertrand, M., J. Pan and E. Kamenica (2013), Gender Identity and Relative Income within Households. *NBER WP #19023*
- 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. 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. Manuscript, Aarhus University.
- Brock, W. A. and S. N. Durlauf (2001), Interactions-Based Models. (vol. 5, pp. 3297-3380) in J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*. Amsterdam: North-Holland.
- 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.
- 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 (2014), Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. Forthcoming in the *Journal of Human Resources*.
- Cunha, F. and J. J. Heckman (2007), The technology of skill formation. *American Economic Review* 97(2): 31-47.
- 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. Forthcoming in the *Quarterly Journal of Economics*.
- Dahl, G., K. V. Løken and M. Mogstad (2014), Peer Effects in Program Participation. *American Economic Review* 104(7): 2049-74.
- Ellingsen, T. and M. Johannesson (2008), Pride and Prejudice: The Human Side of Incentive Theory. *American Economic Review* 98(3): 990-1008.

- 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.
- Goldin, C. and L. F. Katz (2008), *The Race Between Education and Technology*. Harvard University Press. Cambridge, MA.
- 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
- Jalava, N., J. S. Joensen, and E. Pellas (2014), Grades and Rank: Impacts of Non-Financial Incentives on Test Performance. Forthcoming in *Journal of Economic Behavior and Organization*.
- Jewitt, I. (2004), Notes on the shape of distributions. Manuscript, Oxford University.
- Jia, R. and T. Persson (2014), Ethnicity in Children and Mixed Marriages: Theory and Evidence from China. Manuscript, IIES Stockholm University.
- 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 (2014), Mathematics and Gender: Heterogeneity in Causes and Consequences. Forthcoming in the *Economic Journal*.
- 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 for Children and Education (2013), UU-centrenes vejledning (In English: Advice from the Centers for Youth Education). UNI-C, Ministry for Children and 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.
- Powell, A.G., E. Farrar and D. K. Cohen (1985), *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*. Houghton Mifflin: Boston.
- Qureshi, J. (2013), Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital. Manuscript, University of Illinois at Chicago.
- Steele, C. (1997), A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist*, 52: 613-629.
- 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.

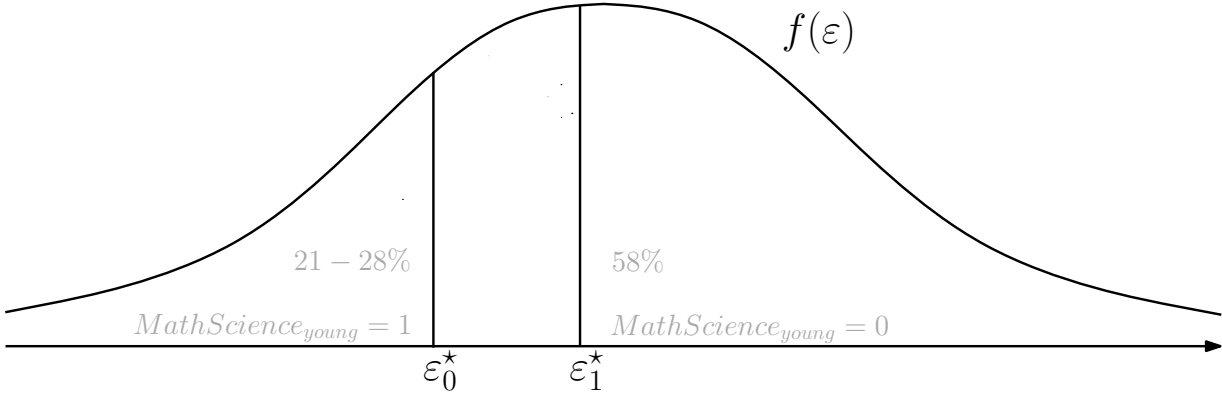
Appendix A. Model Figures

Figure A1. Distribution of Intrinsic Incentives and Math-Science cut-offs.

(a) All



(b) Boys



(c) Girls

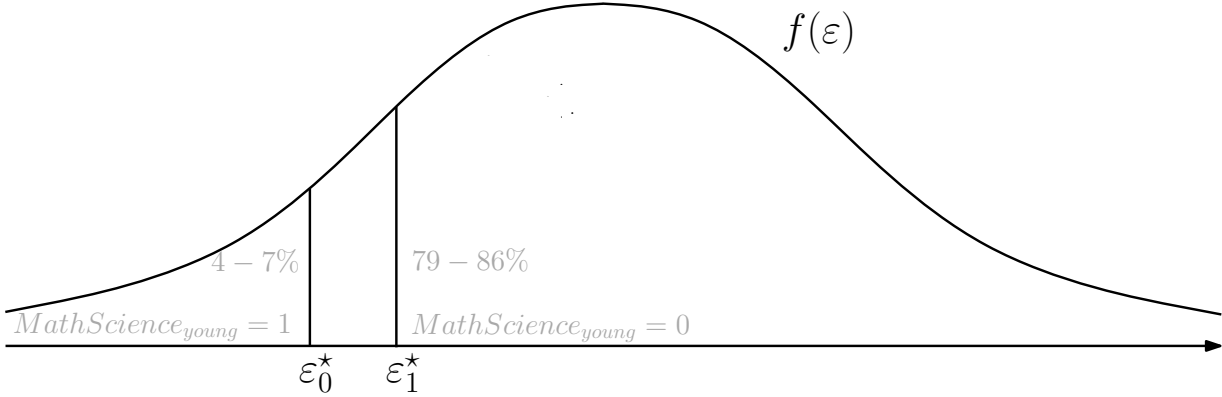
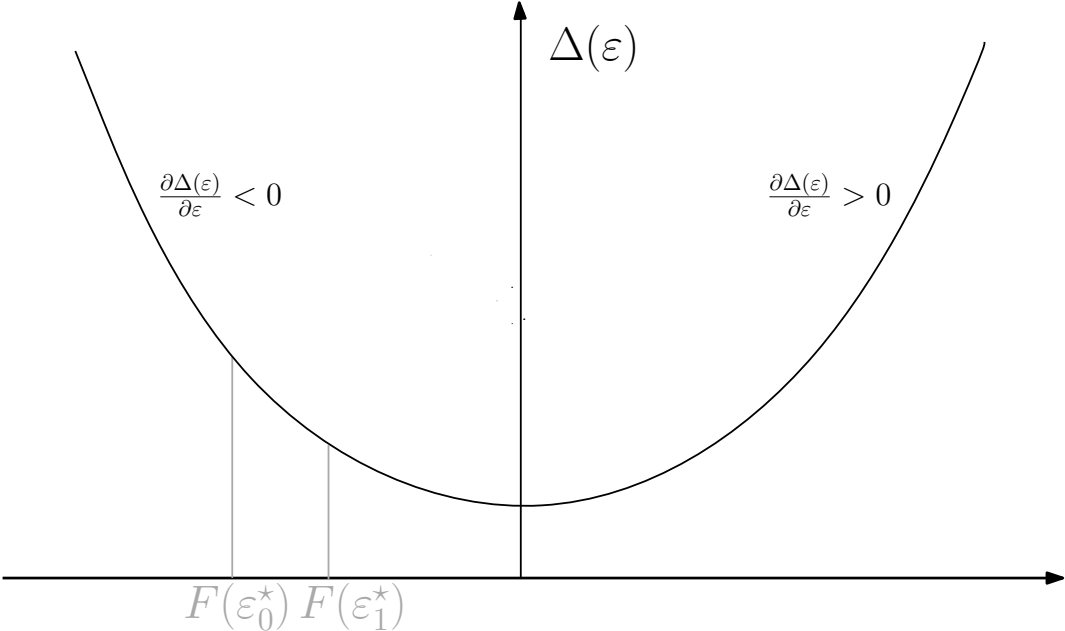


Figure A2. Non-Conformity and Reputational Returns.

(a) All



Note: Figure A2 shows that the reputational return is lowest when choosing Math-Science is the modal choice. It further shows that as we move to the left and Math-Science becomes a more rare choice, then the pride will become even stronger when having an older Math-Science sibling and the prejudice even stronger if having a non-Math-Science older sibling.

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 who unexpectedly introduced the pilot scheme combining advanced Math with advanced Chemistry (PilotSchool=0, PilotIntro=1) for the older sibling, and schools who already had adopted the pilot scheme (PilotSchool=1, PilotIntro=0). The two rows at the bottom summarize information for younger sibling entering high school in the cohorts 1988-97 (age gap ≤ 10 years) and the cohorts 1988-91 (age gap ≤ 4 years).

Table B3. Descriptive Statistics by Course Choice of Older Sibling.

Younger Sibling		Older Sibling		Older Sibling		All	
		MathScience = 0		MathScience = 1			
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. Descriptive Statistics by Pilot School Status of Older Sibling.

Younger Sibling		Older Sibling		Older Sibling		Older Sibling		All	
		Pilot School = 0		Pilot School = 0		Pilot School = 1			
Gender	Course Choice	Pilot Intro = 0		Pilot Intro = 1		Pilot Intro = 0		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 who unexpectedly introduced the pilot scheme combining advanced Math with advanced Chemistry (PilotSchool=0, PilotIntro=1) for the older sibling, and schools who already had adopted the pilot scheme (PilotSchool=1, PilotIntro=0).

Table B5. Descriptive Statistics 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		MeanDiff	Older Sibling Pilot School = 1 Pilot Intro = 0		
	Mean	Std.Dev.	Mean	Std.Dev.		Mean	Std.Dev.	MeanDiff
Male	0.4495	0.4975	0.4636	0.4990	-0.0141	0.4406	0.4966	0.0089
<i>High school cohort</i>								
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	<i>-0.0208</i>	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
Short higher education	0.0356	0.1854	0.0451	0.2076	-0.0095	0.0374	0.1898	-0.0018
Medium higher education	0.2642	0.4409	0.2474	0.4318	0.0168	0.2628	0.4403	0.0014
Long higher education	0.2252	0.4178	0.2116	0.4087	0.0136	0.2604	0.4389	-0.0352
STEM field (narrow)	0.2229	0.4163	0.2370	0.4255	-0.0141	0.2449	0.4301	-0.0220
STEM field (broad)	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
Short higher education	0.0614	0.2401	0.0751	0.2638	-0.0138	0.0736	0.2612	-0.0122
Medium higher education	0.3424	0.4746	0.3306	0.4707	0.0117	0.3393	0.4736	0.0031
Long higher education	0.0726	0.2595	0.0578	0.2335	0.0148	0.0732	0.2606	-0.0006
STEM field (narrow)	0.0471	0.2118	0.0439	0.2051	0.0031	0.0513	0.2206	-0.0042
STEM field (broad)	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
Older sibling 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 bom	0.7930	0.4052	0.8000	0.4002	-0.0070	0.7815	0.4133	0.0114
Second bom	0.1667	0.3728	0.1607	0.3675	0.0060	0.1798	0.3841	-0.0131
Third bom 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
Full siblings	0.9736	0.1605	0.9792	0.1428	-0.0056	0.9837	0.1266	-0.0102
Grew up together	0.9590	0.1983	0.9618	0.1917	-0.0029	0.9699	0.1709	-0.0109

Note: 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 Peer 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 <i>MathScience</i>					
<i>PilotIntro</i>			0.203 *** (0.046) [0.072]	0.226 *** (0.050) [0.073]	0.203 *** (0.051) [0.065]	
	Outcome: Younger Sibling <i>MathScience</i>					
Older Sibling <i>MathScience</i>	0.493 *** (0.034) [0.140]	0.576 *** (0.038) [0.153]	0.572 *** (0.039) [0.151]	1.453 *** (0.350) [0.360]	0.937 *** (0.309) [0.222]	0.828 ** (0.343) [0.195]
Control Variables:						
Entry Cohort Fixed Effects		+	+		+	+
Sibling pair gender composition		+	+		+	+
Sibling pair in same High School			+			+
<i>Parental variables (for mother and father):</i>						
Age, Highest Completed Education, and Income		+	+		+	+
HS Mean of Highest Completed Education and Income			+			+
Mean of Older Sibling <i>MathScience</i>				0.328		
Mean of Younger Sibling <i>MathScience</i>				0.190		
Number of Sibling Pairs				7,786		

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table B7. Estimates of Peer Effects: Cohort Variation

		Parameter Estimates (Standard Errors)					
		OLS			2SLS		
	N	(1)	(2)	(3)	(4)	(5)	(6)
First-stage: Older Sibling MathScience							
<i>PilotIntro</i>					0.080 ***	0.079 ***	0.074 ***
Younger Sibling 1988-90, ≤4y	7,079				(0.018)	(0.017)	(0.017)
<i>PilotIntro</i>					0.068 ***	0.071 ***	0.065 ***
Younger Sibling 1988-91, ≤4y	7,786				(0.017)	(0.016)	(0.017)
<i>PilotIntro</i>					0.069 ***	0.072 ***	0.066 ***
Younger Sibling 1988-92, ≤4y	7,873				(0.017)	(0.016)	(0.017)
<i>PilotIntro</i>					0.070 ***	0.072 ***	0.066 ***
Younger Sibling 1988-93, ≤4y	7,876				(0.017)	(0.016)	(0.017)
Reduced-form: Younger Sibling MathScience							
<i>PilotIntro</i>		0.041 ***	0.037 **	0.028 *			
Younger Sibling 1988-90, ≤4y		(0.015)	(0.015)	(0.015)			
<i>PilotIntro</i>		0.036 **	0.032 **	0.023			
Younger Sibling 1988-91, ≤4y		(0.015)	(0.014)	(0.014)			
<i>PilotIntro</i>		0.038 ***	0.034 **	0.024 *			
Younger Sibling 1988-92, ≤4y		(0.014)	(0.014)	(0.014)			
<i>PilotIntro</i>		0.038 ***	0.034 **	0.024 *			
Younger Sibling 1988-93, ≤4y		(0.014)	(0.014)	(0.014)			
Outcome: Younger Sibling MathScience							
Older Sibling <i>MathScience</i>		0.141 ***	0.148 ***	0.146 ***	0.510 **	0.473 **	0.376 *
Younger Sibling 1988-90, ≤4y		(0.010)	(0.010)	(0.010)	(0.206)	(0.197)	(0.205)
Older Sibling <i>MathScience</i>		0.140 ***	0.147 ***	0.145 ***	0.523 **	0.447 **	0.346
Younger Sibling 1988-91, ≤4y		(0.009)	(0.010)	(0.010)	(0.231)	(0.205)	(0.221)
Older Sibling <i>MathScience</i>		0.139 ***	0.145 ***	0.143 ***	0.540 **	0.465 **	0.367 *
Younger Sibling 1988-92, ≤4y		(0.009)	(0.009)	(0.009)	(0.229)	(0.203)	(0.220)
Older Sibling <i>MathScience</i>		0.139 ***	0.145 ***	0.143 ***	0.541 **	0.465 **	0.367 *
Younger Sibling 1988-93, ≤4y		(0.009)	(0.009)	(0.009)	(0.229)	(0.203)	(0.219)
Control Variables:							
Entry Cohort Fixed Effects			+	+		+	+
Sibling pair gender composition			+	+		+	+
Sibling pair in same High School				+			+
<i>Parental variables (for mother and father):</i>							
Age, Highest Completed Education, and Income			+	+		+	+
HS Mean of Highest Completed Education and Income				+			+

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Table B8. Estimates of Peer Effects: Randomly Matched Sibling Pairs

	Parameter Estimates (Standard Errors)							
	Randomly Matched Only Children				Randomly Matched Only Children (Same High School)			
	OLS		2SLS		OLS		2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	First-stage: Older "Sibling" MathScience							
<i>PilotIntro</i>			0.136 *** (0.021)	0.118 *** (0.030)			0.126 *** (0.022)	0.066 ** (0.031)
	Reduced-form: Younger "Sibling" MathScience							
<i>PilotIntro</i>	-0.002 (0.028)	0.014 (0.028)			0.022 (0.029)	0.001 (0.029)		
	Outcome: Younger "Sibling" MathScience							
Older "Sibling" MathScience	-0.016 (0.019)	-0.007 (0.020)	-0.019 (0.231)	0.115 (0.234)	0.021 (0.020)	0.016 (0.021)	0.357 (0.512)	0.017 (0.433)
Control Variables:								
Entry Cohort Fixed Effects		+		+		+		+
"Sibling" pair gender composition		+		+		+		+
"Sibling" pair in same High School		+		+		+		+
<i>Parental variables (for mother and father):</i>								
HS Mean of Highest Completed Education and Income		+		+		+		+
Mean of Older "Sibling" MathScience			0.265				0.261	
Mean of Younger "Sibling" MathScience			0.190				0.182	
Number of Randomly Matched Sibling Pairs			2,219				1,989	

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively. All 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 the rank of their random number within the 1984-87 and 1988-91 cohorts; e.g. the only child with the lowest random number enrolling in 1984-87 is the matched older placebo sibling to the only child with the lowest random number enrolling in 1988-91. Just as for our main sample, we condition on placebo siblings having an age difference of mostly four years. These results are displayed in columns (1)-(4). Second, the placebo siblings are matched based on the rank of their random number within the same high school as this is a stronger test of whether the schools who were previously pilot schools may have developed an expertise in math and science. These results are displayed in columns (5)-(8).

Table B9. Estimates of Peer Effects: Strength of Ties

	Parameter Estimates (Standard Errors)							
	Older and Younger Sibling							
	Full Siblings		Same Household		Different High School		Same High School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	First-stage: Older Sibling <i>MathScience</i>							
<i>PilotIntro</i>	0.066 *** (0.018)	0.061 *** (0.017)	0.071 *** (0.018)	0.066 *** (0.017)	0.069 * (0.040)	0.065 * (0.039)	0.068 *** (0.019)	0.062 *** (0.018)
Relative to overall first-stage	0.971	0.938	1.044	1.015	1.015	1.000	1.000	0.954
	Reduced-form: Younger Sibling <i>MathScience</i>							
<i>PilotIntro</i>	0.033 ** (0.015)	0.020 (0.014)	0.035 ** (0.015)	0.022 (0.015)	0.082 ** (0.032)	0.060 * (0.033)	0.025 (0.016)	0.015 (0.016)
	Outcome: Younger Sibling <i>MathScience</i>							
Older Sibling <i>MathScience</i>	0.500 ** (0.241)	0.333 (0.238)	0.490 ** (0.224)	0.326 (0.221)	1.196 (0.784)	0.923 (0.689)	0.373 (0.245)	0.251 (0.254)
Control Variables:								
Entry Cohort Fixed Effects		+		+		+		+
Sibling pair gender composition		+		+		+		+
Sibling pair in same High School		+		+		+		+
Parental variables (for mother and father):								
Age, Highest Completed Education, and Income		+		+		+		+
HS Mean of Highest Completed Education and Income		+		+		+		+
Mean of Older Sibling <i>MathScience</i>	0.330		0.331		0.305		0.333	
Mean of Younger Sibling <i>MathScience</i>	0.191		0.191		0.163		0.196	
Number of Sibling Pairs	7,610		7,496		1,439		6,347	

Note: Significance at a 1%, 5%, and 10% level are denoted by ***, ** and *, respectively.

Appendix C. Direct and Indirect Birth-Order Effects

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. This Appendix supplements Section 6.2 by providing details on how we can identify all direct birth order effects without imposing additional parameter restrictions. We exploit unique features of the institutional setup and analyze the educational choices in sibships of three ("triplets").

To illustrate, we simplify the linear projections (10) and (11) to exclude other characteristics:

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

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

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 .³⁸

³⁸ 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.

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 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:

$$\text{MathScience}_1 = \pi_0 + \gamma \text{PilotIntro}$$

$$\text{MathScience}_2 = \beta_{02} + \beta_{12} \text{MathScience}_1 + \gamma \text{PilotIntro}$$

$$\begin{aligned} \text{MathScience}_3 &= \beta_{03} + \beta_{13} \text{MathScience}_1 + \beta_{23} \text{MathScience}_2 \\ &= \beta_{03} + \beta_{13}(\pi_0 + \gamma \text{PilotIntro}) + \beta_{23}(\beta_{02} + \beta_{12}(\pi_0 + \gamma \text{PilotIntro}) + \gamma \text{PilotIntro}) \end{aligned}$$

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:

$$\frac{\partial \text{MathScience}_2}{\partial \text{PilotIntro}} = (\beta_{12} + 1)\gamma = \delta_2 + \gamma = \theta_2$$

$$\frac{\partial \text{MathScience}_3}{\partial \text{PilotIntro}} = (\beta_{23}\beta_{12} + \beta_{23} + \beta_{13})\gamma = \delta_3 + \beta_{23}\gamma = \theta_3$$

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 γ : $\beta_{23} = \frac{\theta_3 - \delta_3}{\gamma}$. 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 exposed to the unexpected 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. This 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: $\beta_{12} = 0.24$, $\beta_{13} = -0.23$, $\beta_{23} = -0.42$, but too noisy to draw strong inference on direct and indirect birth-order effects. 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.