11-2019

Rethinking education choices: The effect of surveys

Luca FACCHINELLO
Singapore Management University, lfacchinello@smu.edu.sg

Juanna Schroter JOENSEN

Gregory Francisco VERAMINDI

Rethinking Education Choices: The Effect of Surveys

Luca Facchinello*
Juanna Schrøter Joensen‡
Gregory Francisco Veramendi§

November 20, 2019

latest version here

Abstract: Can surveys affect human capital investments? This paper examines whether individual education choices and outcomes are affected by a survey posing questions related to investments, performance, preferences, and expectations. We have administrative data for the whole Swedish population to which an extensive education survey was administered to randomly drawn samples of primary school students and their parents. This constitutes a large-scale randomized social experiment for testing whether responding to survey questions alters behavior. We observe complete education and labor market histories until the individuals are 31-36 years old. The causal effect of the survey on both short- and long-term outcomes is generally not significantly different from zero. We find, however, that being surveyed affects educational attainment, earnings, and employment in the early career for some subgroups. We assess heterogeneity in estimated causal effects and sibling spillovers in order to get at potential mechanisms. The patterns are consistent with the importance of increased awareness (or salience) and short-lived information "nudges".

JEL: C42, C83, I21, I24, J24.
Keywords: Survey, Awareness, Information, Education Choices.

*Singapore Management University. Email: lfacchinello@smu.edu.sg
‡University of Chicago, Stockholm School of Economics, TrygFonden's Centre for Child Research–Aarhus University, and IZA. Email: jjoensen@uchicago.edu
§University of Munich (LMU). Email: gregory.veramendi@econ.lmu.de.
We thank seminar participants at the Stockholm School of Economics and the University of Chicago, and participants at the 22nd Annual Meetings of the Society of Labor Economist (SOLE) and the 39th Association for Public Policy Analysis & Management (APPAM) Annual Conference for valuable feedback. We appreciate insightful discussions by Krzysztof Karbownik and Björn Öckert. We gratefully acknowledge financial support from the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond) grant P12-0968 enabling the data collection for this project. We are also grateful to the Stockholm School of Economics for hosting our research project. The usual disclaimers apply.

1 Introduction

Empirical work in the social and health sciences depends on the use of survey data. Surveying individuals may draw their attention to risks, returns, or choices previously not salient to them. This may change their subsequent behavior. If so, it may bias parameter estimates and any conclusions drawn from survey data. In this paper, we assess whether surveys can causally affect educational investments. We also assess for *whom* and *how* surveys can change education choices and outcomes.

Does responding to a survey designed to evaluate own performance in school and reflect on future education opportunities cause individuals to rethink their education choices? We answer this question by linking survey data to comprehensive administrative data for the whole Swedish population. The extensive education survey was administered to a randomly drawn sample of two cohorts of 3rd graders in the late 80s and early 90s. The survey thus constitutes a randomized social experiment for testing whether reflecting on survey questions alters behavior. We observe education and labor market outcomes until individuals are 31-36 years old. Importantly, these are from administrative registers and *not* reported by the individuals themselves. This means that (i) measurement error is minimal in our data, (ii) we can conduct balancing tests on a rich set of pre-determined characteristics of surveyed and non-surveyed individuals to corroborate random assignment, and (iii) we can analyze the impact of the survey on both shorter- and longer-term outcomes.

How can responding to an extensive education survey change investment in education? The "treatment" of being surveyed is a bundle of different types of information. This means that there are several potential channels through which the survey can affect education choices. Four potential channels are: First, the children need to evaluate themselves and their abilities – also relative to their peers' abilities. They also take multiple cognitive aptitude tests. Even if they are not informed of their test scores, the test situation may still convey information to those in the tails; e.g. if they could not reach the end of the test, were not able to answer many test items, or finished before their

peers and confidently solved all test items.[1] Second, the survey required them to state their preferences; e.g. state their future education plans and desired occupation. This can be seen as a "nudge" to rethink and evaluate goals, means, and returns in a forward-looking manner.[2] This could lead to more well-considered choices and less "mismatch" between individual abilities, education, and career choices.[3] Third, the survey could increase awareness or the salience of choices not previously considered. The connection between early academic choices and the tracked schooling system may not be clear to the child and the parents.[4] There may also be an information asymmetry in that parents with higher education may be better informed than parents who dropped out of school after compulsory schooling. The survey asks for reflection on these education choices, including continuing to high school and college. This information could affect choices by increasing awareness of choices not previously considered; especially for parents who never themselves took these education paths. If so, this has substantive implications for how to model education choices. Limited attention models, where individuals simplify complex choice settings by only considering a limited subset of choices, have existed at least since Simon (1955).[5] However, such models are not yet widespread in the context

[1]Several studies find considerable uncertainty about abilities and room for learning through grades and other feedback in education (Altonji, 1993; Arcidiacono, 2004; Arcidiacono et al., 2011, 2012; Zafar, 2011; Facchinello, 2015). Our paper is also related to the literature estimating the effects of providing performance feedback (Wagner, 2015), top-3 performance feedback (Bursztyn, 2017; Jalava et al., 2015), relative performance feedback (Azmat, 2010; Tran, 2012; Martinez, 2014; Azmat et al., 2016; Fischer and Wagner, 2017; Davis et al., 2017), and relative versus absolute grading (Jalava et al., 2015; Czibor et al., 2014) in education settings.

[2]In the survey wave after compulsory schooling completion (12th grade) individuals are also asked how they made their education choices of elective courses and how they decided to enroll in high school or not. This is also "nudging" them to evaluate the optimality of their past choices. Some who may otherwise have dropped out after compulsory schooling may thus decide to enroll in high school after having stopped out of school for a year or two. (Quasi-)Experiments providing information to make individuals reflect on the optimality of current behavior have found beneficial effects (Bandiera, 2017; Bobba and Frisancho, 2016; Pistolesi, 2017).

[3]Surveys may affect education choices by providing information to rethink education choices and potentially change expectations. Our paper is therefore also related to the literature analyzing the short-term effects on educational attainment of providing statistical information on the population distribution of the general returns to education (Nguyen, 2008; Jensen, 2010; Carrell, 2017) or on college major-specific characteristics (Wiswall and Zafar, 2015; Hastings et al., 2015; Pekkala Kerr et al., 2015). The empirical results in this literature are mixed – ranging from no to positive effects for some subgroups.

[4]More information on education tracks may also improve outcomes in terms of more sorting on test scores across high school tracks and less dropout (Goux et al., 2014).

[5]See e.g. Barberis and Thaler (2003) and DellaVigna (2009) for reviews on contexts in finance and economics where limited attention has been found important.

of human capital accumulation and education choice.[6] Fourth, parents also respond to questions about school inputs, school choices, and how much of their time they devote to their children – in particular to their schooling investments. This could lead parents to invest more time in their children's skill accumulation and schooling.[7]

We shed light on the potential channels by assessing heterogeneity in treatment effects estimated under different identifying assumptions (a within-municipality and a between-school estimators) across subgroups with different levels of parental education.

The causal effect of being surveyed on both short- and long-term outcomes is generally not significantly different from zero. We find, however, that being surveyed changes grades, educational attainment, income, and job stability in the early career for some sub-populations. If anything, being surveyed increases inequality as those with low parental education tend to get lower grades, education, and income, while those with high parental education tend to select into better "matched" and more stable careers if surveyed.

We try to disentangle the mechanisms by examining whether it makes a difference *when* and *who* is surveyed – the children themselves or also their parents. We merge data on siblings in order to estimate potential information spillover effects. Siblings who were not directly affected by the survey, could only be affected if there are information spillovers through social interaction with their siblings and parents – a potentially important determinant of education choice (Joensen and Nielsen, 2018; Dustan, 2018; Qureshi, 2018). We exploit that the parent is surveyed when their non-surveyed children (i.e. the siblings of the surveyed child) are at different stages of their educational paths. Some older siblings are about to make critical decisions on whether to enroll in high school or in college. Therefore, the importance of education may become particularly salient for these siblings as the parent reflects on education and career choices. We find some evidence

[6] See e.g. the recent surveys by Koch et al. (2015), Lavecchia et al. (2016), and Damgaard and Nielsen (2018) for the potential importance of limited awareness more generally in the context of education choice.

[7] Providing parents with information about school quality (Hastings and Weinstein, 2008), the benefits of taking math and science classes (Harackiewicz et al., 2008), and academic performance (Dizon-Ross, 2019) is found to change investments child education. More parental involvement in child's schooling is found to improve their child's, and even their child's peers', school attendance (Avvisati et al., 2014). Cunha et al. (2010) find that measured parental investments account for 15% of the variation in educational attainment. Heckman and Mosso (2014) provide a recent comprehensive review of this emerging literature.

that siblings who are at critical junctions when the parent survey is administered are "nudged" into better "matched" careers. Timing therefore seems to be crucial.

If being surveyed changes behavior, then it has implications for both the external and internal validity of studies based on survey data. The total "treatment" effect of being surveyed is thus interesting *per se*.[8] The empirical issues of dealing with measurement error and validation have been extensively studied.[9] The fact that repeated surveying can alter individual survey response patterns is also well-established. For example, related to the monthly Current Population Survey (CPS) rotation group bias and its impact on unemployment estimates (Bailar, 1975; Solon, 1986).

Our paper is also related to the literature on "*Hawthorne effects*" first reported following an extensive research programme investigating methods of increasing productivity in the Western Electrical Company's Hawthorne Works in Chicago during the 1920s and 30s (Roethlisberger and Dickson, 1939; Franke and Kaul, 1978). No matter what change was introduced to working conditions, the result was increased worker productivity. The Hawthorne effect was originally defined as "*an increase in worker productivity produced by the psychological stimulus of being singled out and made to feel important*". Subsequently the original findings have been questioned (Parsons, 1974; Adair, 1984; Jones, 1992) and the definition has been broadened to also refer to treatment response and potential outcomes rather than productivity per se. The Hawthorne effect has implications for the external validity of randomized controlled trials (RCTs) more generally if there is a demonstrable benefit from participating in a trial. Most RCT designs are unable to quantify the magnitude of the Hawthorne effect because its defining features (e.g. extra attention by researchers and higher levels of clinical surveillance) apply equally to treatment and control groups. In the domain of health, McCarney et al. (2007) find that more intensive follow-up of individuals in a placebo-controlled clinical trial resulted in better outcomes than minimal follow-up.

[8]Most countries administer Household and Labor Force Surveys (LFS). Much of the research on education, labor market, household finance, health, and retirement choices is based on survey data. For example, the Panel Study of Income Dynamics (PSID) and the National Longitudinal Survey of Youth 1979 (NLSY79) and 1997 (NLSY97). The construction of these panel surveys has been crucial for many of the recent methodological advances in microeconometrics and applied microeconomics more generally.

[9]See e.g. Bound et al. (2001) for an extensive survey of the literature.

However, whether responding to surveys can actually alter behavior is still an open question. The only study we know of (Zwane et al., 2011) finds mixed evidence. Zwane et al. (2011) randomly assign individuals in five field experiments – three on health and two on micro-lending – to respond to survey questions on health and/or household finance. They find that responding to health-related questions significantly alters health-related behavior. Those randomized to take the health survey had significantly higher take-up of medical insurance and increased use of water treatment products. They also find that this leads to biased estimates of improved water source quality – despite random assignment of water source quality. These results indicate that researchers should be cautious when administering extensive and repeated surveys, since they may alter the estimated treatment effects of those surveyed by changing their behavior. However, these results seem to be context-dependent as they detect significant survey effects in the domain of health, but not in the domain of micro-lending.

The contribution of this paper to the literature is sixfold: First, we merge the random survey sample to administrative data for the whole Swedish population. Our sample size is therefore much larger and the measurement error in education choices and outcomes is minimal. Second, we analyze the effect of being surveyed in a developed (instead of developing) country. Third, we examine the domain of education (instead of health and micro-lending). This is important if impacts are context-specific. For example, if survey effects only arise in settings where individuals previously ignored some of their potential choices and opportunities; e.g. enrolling in high school or college. In such settings, the survey can make these choices more salient, divert the focus to rethink priors, and spur individuals to take otherwise missed opportunities; e.g. more rewarding education paths. Fourth, we have access to complete medical birth records and a range of measures of family composition and resources to conduct balancing tests of pre-determined characteristics of surveyed and non-surveyed individuals. Fifth, we observe individuals for 18-21 years after random survey assignment. This allows us to analyze both shorter- and longer-term outcomes, which is important if the strength of the effect of being surveyed diminishes or amplifies over time. Lastly, we link the focal individuals to their siblings and analyze

spillovers to shed light on potential channels through which surveys affect education and labor market outcomes.

The rest of the paper proceeds as follows: the next section spells out the details of the institutional setting, survey sampling scheme, and how we exploit these in our empirical strategy. Section 3 presents the data, descriptive statistics, and balancing tests. Section 4 presents the empirical results, while Section 5 concludes.

2 Institutional Setting and Empirical Strategy

This section first provides some background on the institutional setting and the Swedish schooling system the surveyed individuals and their cohorts were facing. Second, we describe the sampling scheme of the Evaluation Through Follow-up (ETF) cohort-sequential longitudinal survey. Third, we describe the empirical strategy we use to identify the effect of surveys on education choice.

2.1 Survey Sampling Scheme

The ETF survey was administered by the Department of Education and Special Education, Gothenburg University, in collaboration with Statistic Sweden.[10] The survey was constructed through a multistage sampling scheme with stratification at the municipal level: (1) systematically draw two municipalities at random per stratum (13 strata) plus the three largest municipalities (Stockholm, Gothenburg, and Malmö), (2) randomly draw classes within sampled municipalities, and (3) survey all students in each sampled class.

More specifically, the survey sample selection design was as follows. In stage (1) the three largest municipalities (Stockholm, Gothenburg, and Malmö) were selected with probability one, while the remaining 281[11] Swedish municipalities were categorized into 13 strata according to their population (above or below 25,000 inhabitants), proportion of "socialist" mandates (above or below 50%), the share employed in public administration

[10] Härnqvist (1998) provides additional details on the construction of the survey.
[11] 283 in the last wave.

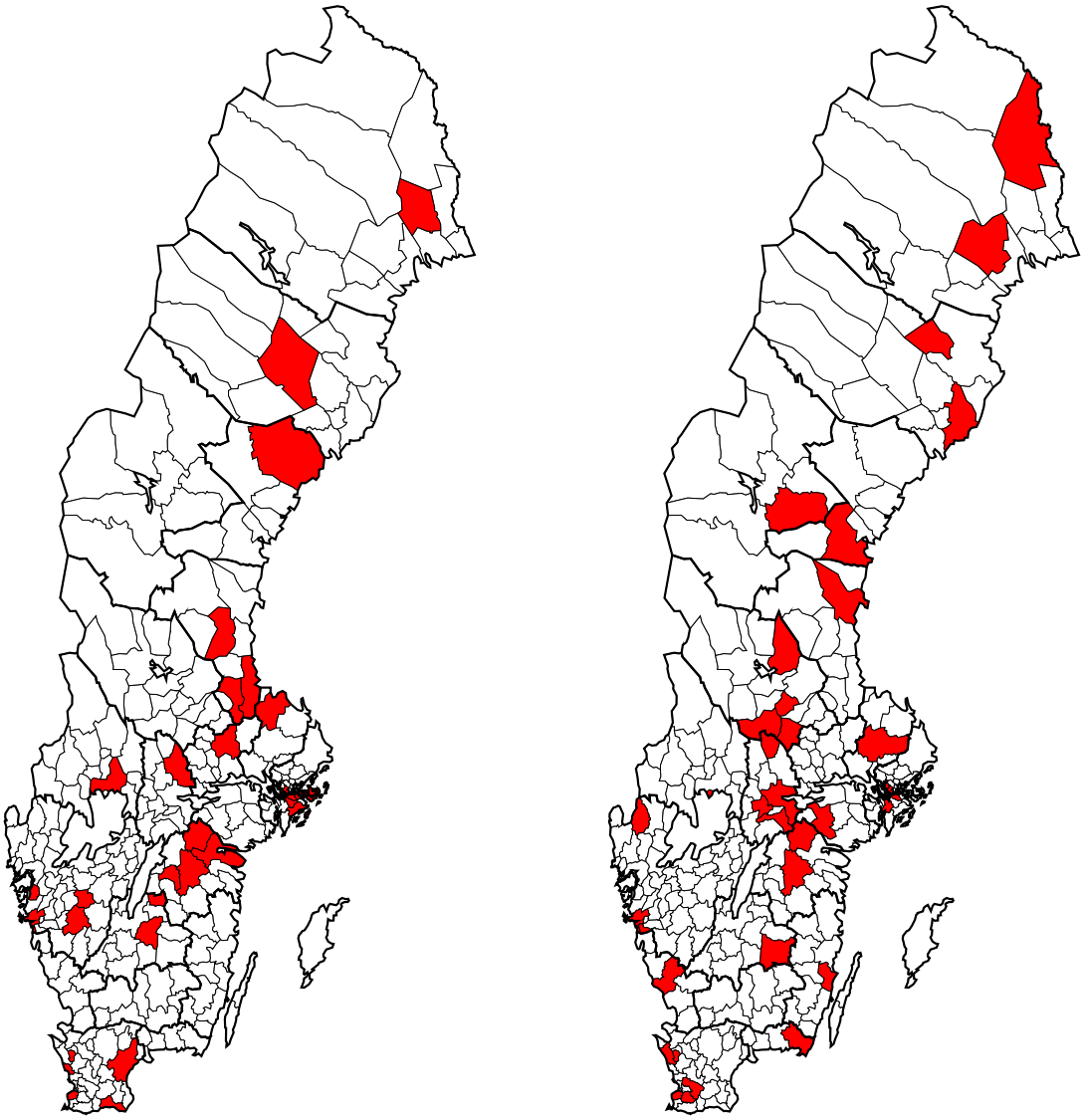
(above or below 25%), and the proportion of immigrant pupils (above or below 8%).[12] Within each stratum, the municipalities were assigned a sampling probability weight, p_m, proportional to their share of pupils in the relevant school cohort. Finally, two municipalities were sampled at random (conditional on p_m) from each stratum. Thus, a total of 29 municipalities were sampled and larger municipalities were more likely to be sampled. Figure 1 displays a map of Sweden with each of the sampled municipalities in stage (1) for each cohort.

In stage (2) Statistic Sweden's (SCB) class register was used to randomly sample classes: 3rd graders in the school-years 1986/87(ETF77) and 1991/92 (ETF82). Unfortunately, these class registers have not been kept in SCB's archives. This means that we can not perfectly measure "treatment" and "control" group assignment, since we can not perfectly measure who is in the "control" group: i.e. those who attended 3rd grade in the sampled municipalities in the same year, but were not in a sampled class. The sample selection criteria for the classes were that: (2.i) 17 classes in each sampled municipality were selected at random, but with the two exceptions that (2.ii) all classes in small municipalities with 17 or fewer classes were sampled and (2.iii) 29 classes were selected at random in the three largest municipalities (Stockholm, Gothenburg, and Malmö). The class sampling was made in January in the relevant school-year (1986/87 and 1991/92).

In stage (3), all students registered in the sampled classes were surveyed on April 15 in the relevant school-year, hence "treated". The first cohort (ETF77) was surveyed in 3rd, 6th, and 12th grade – three years after compulsory schooling and stipulated academic high school graduation – whereas the last cohort (ETF82) was only surveyed in 6th and 12th grade.[13] The students' parents were also surveyed once when the child was in 6th grade. Table 1 provides an overview of the administered surveys and tests for each cohort.

[12]The exact procedure of selecting municipalities is extensively documented in Emanuelsson (1979).

[13]Note that randomization also happened in 3rd grade for the ETF82 cohort despite them not being surveyed until 6th grade.



(a) 1977 cohort (b) 1982 cohort

Figure 1: Sampled municipalities

Note: The Figure displays a map of Sweden with the sampled municipalities in each ETF-cohort 1977 and 1982, respectively, shaded in red.

2.2 Control and Treatment Group Assignment

In this section, we explain how we make the best possible approximation to the "control" group conditional on the available data.

First, we select those who were in 9th grade in 1992/93 and 1997/98. The two main drawbacks with this selection method is that it assumes no grade retention and no students selecting in and out of schools during the six-year period from grade 3 to grade 9. Students who may re-take or skip a grade will be misclassified according to

Table 1: Treatment Assignment Overview

	Parent survey		Child survey			Aptitude test	
	3rd	6th	3rd	6th	12th	3rd	6th
1977 cohort		T	T	T	T	T	T
1982 cohort		T		T	T	T	

The Table displays an overview of the variation in treatment assignment over grades for each of the three ETF-cohorts 1977 and 82.

this assumption. However, grade retention was extremely rare for these cohorts. We exclude those who immigrated to Sweden after January 1 in the academic year they attended grade 3, since those assigned the survey (i.e. the "treatment" group) will by construction only include those who were registered in a Swedish school at the time of random assignment. We also have information on the municipality at birth as well as the municipality of residency of the parents in December 1990 (when the ETF77 cohort attended 7th grade) and in December 1991 (when the ETF82 cohort attended 3rd grade and was sampled). This allows us to almost perfectly measure who was in the "control" group in the ETF82 cohort and more accurately measure who was in the "control" group in the ETF77 cohort.

We discuss additional sample selection criteria and refinements to the approximation of treatment and control groups in Sections 2.3 and 3.

2.3 Empirical Strategy

We now turn to describing the empirical strategies we use to identify the effect of surveys on education choice. Particularly, how we exploit the survey sampling scheme to identify the causal effect of being surveyed. This section highlights how – despite successful randomization – we need to consider non-random sorting across units of randomization.

First, we simply compare the outcomes of those surveyed and those non-surveyed within each municipality. This is given by the linear regression:

$$Y_{ismc} = \delta_0 + \delta_1 Survey_{ismc} + \gamma_m + \varepsilon_{ismc} \tag{1}$$

where Y_{ismc} is the outcome for individual i in school s in cohort c and municipality m, *Survey* is an indicator for whether the individual was surveyed, and γ_m is a municipality fixed effect. Given the survey sampling scheme (cf. Section 2.1), δ_1 can be interpreted as a causal effect of being surveyed if (a) individuals, classes and schools are randomly sampled and (b) there is full compliance as everyone assigned to the survey responded to the survey. In other words, if individuals are neither assigned to sampled classes nor refusing to respond to the survey based on the unobservables, ε_{ismc}, in (1) which affect the outcome of interest.

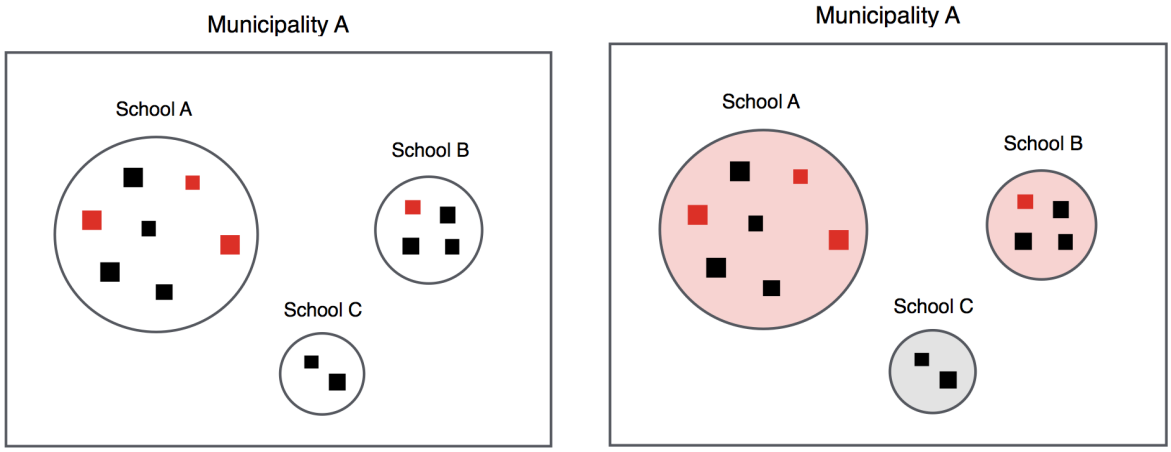
Second, we use a between-school estimator to deal with potential non-random sampling across schools as well as non-random individual non-response. The between-school strategy also assumes (a) random sampling of schools (and classes), but tries to get at potential violations of (b) non-random student non-response by measuring "treatment" at the school level. However, the benefit of not having to assume (b) comes at the cost of the estimated treatment effect being attenuated towards zero.

The following two sub-sections are devoted to providing more details on the within-municipality and between-school estimators we apply. Overall, the empirical strategies trade-off precision and bias in different ways by imposing different identifying assumptions and measuring "survey treatment" at different unit levels. We will discuss the threats to interpreting each of these "survey effects" as causal in even more detail when presenting the data and empirical results in Sections 3 and 4.

2.3.1 Within-Municipality

Figure 2 (a) illustrates how the within-municipality specification (1) simply compares the four treated classes in *Municipality A* with the nine control classes in *Municipality A*. δ_1 can be interpreted as the causal effect of being surveyed if treated classes are not selected based on unobservables, ε_{ismc}, in (1) that affect the educational outcome. This seems reasonable based on the class selection criteria outlined in Section 2.1. However, there are a few empirical issues we need to deal with. First, we only partially observe class assignment in 3rd grade as it is only observed for those who are assigned to treatment

and comply. Thus we need to impute 3rd grade school assignment by survey response and 9th grade school.[14] Second, there may be attrition due to some students moving after randomization occurred. Third, some students may also have been in a different class, school, municipality, or abroad at the time of randomization but otherwise followed the sampled class. Fourth, many students switch schools between 3rd and 9th grade simply because some schools specialize in either younger or older grades. Fifth, some students (more realistically their parents) might have opted out of the survey for privacy reasons. Therefore, both student non-response and mobility between 3rd and 9th grade pose threats to the identification of δ_1. In Section 3.2.1, we assess the credibility of the identifying assumptions by testing for balance on a range of variables determined pre-treatment. When possible, we also try to control for location in 3rd grade and whether the student is foreign born in (1) – which seems to be a good proxy for mobility.



(a) Within-Municipality (b) Between-School

Figure 2: Variation in Treatment Assignment

Note: The Figure displays the two types of exogenous variation we exploit in the data. The left panel (a) displays how we compare treated and control (light shading) units within each municipality when employing the within-municipality identification strategy. The right panel (b) displays how we compare schools who have some classes assigned to treatment and no classes assigned to treatment, respectively, in the between-school identification strategy.

[14]Section 3.1.1 provides more detail on the imputation of schools.

2.3.2 Between-Schools

To deal with some of the potential threats to identification of a causal survey effect using the within-municipality estimator, we also use a between-school estimator. To this end, we compare the outcomes of individuals who were in *ETF schools* (i.e. schools with at least one treated class) to those schools who did not have any classes assigned to treatment. Figure 2 (b) illustrates this identification strategy. *Municipality A* has three schools of which *School A* and *School B* are ETF schools (marked with light shading) and *School C* is a non-ETF school. The between-school strategy essentially compares the average outcomes of the students in the two ETF schools (*School A* and *School B*) to the students in the non-ETF school (*School C*). The between-school estimates are given by the linear regression:

$$Y_{ismc} = \alpha_0 + \alpha_1 ETFschool_{smc} + \gamma_m + \varepsilon_{ismc} \tag{2}$$

where $ETFschool_{smc}$ is an indicator for whether school s in municipality m is an ETF school for individuals in cohort c. α_1 can be interpreted as the causal effect of being in a school where some students were surveyed. Causal inference naturally rests on the assumption that schools are not sampled based on student unobservables, ε_{ismc}, in (2).

The main advantage of this between-school estimator is that there is no need to impose assumptions on student survey non-response. There are two main drawbacks. First, as classes were systematically sampled larger schools are overrepresented within municipality, due to the fact that they have more classes. This problem is particularly severe in small municipalities, where the majority of classes were sampled. This induces a mechanical bias, as small special schools are less likely to be assigned to treatment. This bias can be corrected controlling for school size. Second, some schools – particularly larger schools – have few surveyed students. This means that the estimated "survey effect" will be attenuated towards zero, since the average outcome in "surveyed" ETF schools is an average over a few surveyed students and many non-surveyed students. To accommodate this, we also estimate a specification of (2) where we replace $ETFschool_{smc}$ with the

fraction of students in the school who were assigned to the survey, $FractionSurveyed_{smc}$. This specification captures the intensity of treatment in each school, which minimizes the attenuation bias. However, $FractionSurveyed_{smc}$ may be correlated with the number of students not responding to the survey. To deal with this potential issue, we rather include the predicted fraction of students surveyed in the school, $\widehat{FractionSurveyed}_{smc}$, in (2). We predict the number of classes sampled in each school by: First, using Maimonides' rule (with a 30 students per class cap) for 9th grade enrollment to predict the average class size.[15] Given the number of students surveyed, we approximate how many classes were sampled in each school. There were very few cases in which whole classes did not comply to the survey assignment. In theory, this specification of (2) should minimize selective non-response bias while also providing a reasonable measure of the intensity of treatment at the school level. Empirically, however, it does not provide us any additional insights.

Finally, it should be noted that we assume no spillover effects on untreated units (i.e. individuals, classes, and schools) when making causal inference throughout the paper. That is, we impose the stable unit treatment value assumption (SUTVA) of any unit's outcome being unaffected by another unit's treatment assignment. If this assumption does *not* hold because there are spillovers from those surveyed to their non-surveyed peers in the same school, then both the simple within-municipality (δ_1) and the within-school estimate (β_1) of the "survey effect" may be attenuated towards zero, while the between-school estimate (α_1) will encompass these spillovers. We can get a sense of the presence and strength of potential spillovers by comparing the different estimates. If there are strong spillovers on peers *within* municipalities, then the "control classes" in the within-municipality strategy are partially "treated". This means that the estimated "survey effect" will be attenuated towards zero as the average outcome of those in surveyed classes is compared to the average outcomes of those not surveyed, but potentially affected by

[15] Following Angrist and Lavy (1999), several papers have used Maimonides' rule to estimate the effect of class size on student achievement. Fredriksson et al. (2013) and Fredriksson et al. (2016) exploit the ETF survey data merged with administrative data to estimate the long-term effects of class size and shed light on the underlying mechanisms. Figure D.1 illustrates the relationship between school size and class size as predicted by Maimonides' rule in our sample.

their surveyed peers. For the between-school estimates, these peer spillovers would work in the opposite direction by attenuating the estimated "survey effect" less towards zero, since they compare the average outcomes in schools where some classes are assigned to the survey treatment to the average outcomes in schools where no students are assigned to treatment.

3 Data

In this section, we describe the data and the sample selection.

We merge the ETF survey data to several administrative registers via the unique Swedish individual identifier. Our measures of education choices and outcomes originate from several registers administered by Statistics Sweden (*SCB*): the 9th grade registry (incl. grades in individual courses), the High School registry (incl. grades in individual courses, grade point average (GPA), track and specialization choices), and the Higher Education registry (incl. detailed educational codes for all enrollment spells, course credits accumulated during enrollment, and acquired degrees).

The Multigeneration registry allows us to link children to their parents. It also contains information on family size and composition. Additional background variables are obtained from the longitudinal integration database for health insurance and labour market studies (*LISA*) from which we have yearly observations during the period 1990-2013. The parental background variables we observe include age, civil status, highest completed education, employment, earnings, and disposable family income. We supplement this with earnings information from the Register Based Labor Market Statistics (*RAMS*) for the years 1986-89 and information on disposable family income from the Income and Tax registry (*IoT*) for the years 1978-89. This means that we can measure disposable family income (parental earnings) from birth (age 4) to age 31 for the youngest cohort and from age 1 (age 9) to 36 for the oldest cohort in our sample.

3.1 Sample Selection

We focus on the cohorts who completed compulsory schooling (9th grade) in the school-years 1992/93 and 1997/98. For children who followed the stipulated education path, this corresponds to the cohorts born in 1977 and 1982. 95.48% and 95.18% of students stipulated to graduate from compulsory school grade 9 in 1993 and 1998 were born in 1977 and 1982, respectively. For each birth cohort, we also focus exclusively on the sampled municipalities.

For the samples used in the within- and between-school empirical strategies, we further exclude those attending very small schools in 9th grade; i.e. schools with a graduating cohort of 20 or fewer students. It is difficult to make a reasonable ETF school imputation for these small schools, since we do not know whether a few surveyed students switched to the school or were part of a very small sampled class. This restriction only drops 0.7% and 1.78% of the sample for the ETF77 and ETF82 cohort, respectively.

3.1.1 ETF School Imputation

We observe which school everyone attends at the end of compulsory schooling; i.e. in 9th grade. However, we only observe earlier class and school choices for those surveyed. In order to impute school status for the whole sample – which we need for the within and between-school strategies – we need to approximate how each observed 9th grade school corresponds to each 3rd grade school. To this end, we use school codes in 3rd grade for those assigned to the survey. These codes are linked to the school codes we observe for everyone in 9th grade. The information on (class) school codes thus reveals how many surveyed students in the same 3rd grade (class) school are also in the same 9th grade school. The details of the ETF school imputations are as follows:

Our imputation procedure is as follows: (i) For each grade 9 school code, find the modal grade 3 school code. (ii) Sum the number of students with the modal school code, N_s^3. (iii) Divide N_s^3 by the total surveyed students in the 9th grade school, s. For each grade 9 school code, this yields the fraction of students with the modal grade 3 school code. (iv) Change school status if fewer than a third of students are coming from the

modal grade 3 school. This changes assignment status for one school in the ETF77 cohort and 17 schools in the ETF82 cohort. Note that the number of schools increases over time, but this does not affect school assignments much.

Once control group students are matched up, we calculate the fraction of surveyed students in each grade 9 school. We classify a school to be a non-ETF school if: there are five or fewer surveyed students in the grade 9 school *and* they represent at most 5% of the school. This changes status for 62 and 35 schools in the ETF77 and ETF82 cohort, respectively.

3.2 Descriptive Statistics

Table 2 shows that survey response was high – initially 95% in the 1977 cohort – but there was also some attrition as response rates fell by about 20 percentage points over the nine year period from 3rd to 12th grade. The response rate is generally lower for the 1982 cohort. Teachers administered the survey in school for the 3rd and 6th graders, while the post-compulsory schooling survey was mailed to the child's home address.[16]

Table 2: Percentage of Students and Parents Responding to Survey

Survey	Child			Parent	
Grade	3	6	12	3	6
Cohort 1977	95	91	73	–	77
Cohort 1982	–	87	62	–	75

Table 3 displays the number of students and schools in the sampled municipalities. Around 60% of schools were sampled in the 82 cohort, while only 51% of schools were sampled in the 77 cohort. The fraction of students sampled within each sampled school is similar 33% in the 82 cohort and 29% in the 77 cohort. Tables D.1 to D.6 in Appendix D.2 show averages and standard deviations of the outcome variables. We separately display the variation within and between municipalities and schools, respectively. This reveals

[16]See ETF database for more details and data documentation.

that there is much more variation within municipalities than between for both cohorts – 6 to 22 times. There is also much more variation within schools than between schools for the 1977 cohort, while the difference is smaller for the 1982 cohort.[17] Overall, the tables show that there should be enough variation in survey assignment within municipalities to estimate (1) and between schools to estimate (2).

Table 3: Students and Schools Sampled - Full Sample

	Students	Schools	Sampled Schools	Fraction in Sampled Schools	Fraction Surveyed in Sampled Schools
Cohort 1977	21689	222	109	0.56	0.29
Cohort 1982	27621	341	204	0.73	0.32

3.2.1 Balancing Tests

To corroborate randomness of the sampling scheme, we perform a number of balancing tests on the pre-determined characteristics of the "treatment" and "control" groups.

Tables A.1 to A.11 in Appendix A display the balancing tests for each of the cohorts, under different sample restrictions, and for each of the empirical strategies. Each table displays control group means in the first column. We also present three sets of balancing tests. First, regression tests without and with controls. Second, standardized difference, the difference between the treatment group mean and the control group mean of each observed characteristic, X, scaled by the pooled variance.

The balancing tests indicate that assumption (b) in Section 2.3 of random survey non-response is more of an empirical issue – especially for the 82 cohort – as there are some systematic and significant differences in observed characteristics between "treatment" and "control" groups. Particularly, having divorced parents and being foreign born are predictive of group assignment. Students with disrupted families might have been more likely to drop out of the survey. Being foreign born presumably introduces issues of selective mobility affecting the within-school "survey effect" estimate. To get a sense of the

[17] This is most likely a result of the school choice reforms enacted in the early 1990s; see Bjorklund et al. (2005) for details.

importance of selective mobility may be, we also control for location in 3rd grade (when possible) and whether the student is foreign born in (**??**). 93% and 95% of the students in the 77 and 82 cohorts, respectively, remain in one of the sampled municipalities. Thus mobility should not have a major impact on the analysis. Nevertheless, we also perform the analysis on a restricted sample of Swedish born students. This improves balance on pre-determined characteristics.

Table 4: Summary of Balancing Tests

Identification	**Restriction**	**1977**	**1982**	**All**
Within municipality	-	1*	6*	3*
		0**	3**	2**
		0***	2***	1***
Within municipality	Swedish born	**0***	3*	3*
		0**	2**	1**
		0***	0***	0***
Between school	-	2*	3*	**0***
		0**	1**	**0****
		0***	0***	**0*****

The Table displays an overview of the balancing tests presented in Appendix A. The displayed number indicates how many times the null hypothesis of no mean difference between the treatment and the control group is rejected. *, **, *** indicates how many tests are rejected with a 10%, 5%, and 1% level of significance, respectively. Total number of tests in Tables A.1-A.11: 24. Preferred specifications for each cohort in **blue**.

The overall balance is summarized in Table 4. In particular, when restricting the sample to Swedish born students, there are no statistically significant differences between surveyed and non-surveyed students within municipality for the 1977 cohort (see third column in Tables A.1 and A.2). Controlling for school size, there are no statistically significant differences between students in sampled and non-sampled schools for the pooled sample of the 1977 and 1982 cohort (see second column in Tables A.11 and A.12). Overall, the within-municipality strategy seems most credible for the 1977 cohort, while the between-school strategy seems most credible for the pooled sample as non-random survey non-response seems to have been more of an issue for the 1982 cohort.

3.2.2 Missing Variables

We now turn to analyzing the fundamental issue of non-random survey non-response to assess how it may affect our estimates. First, we analyze to what extent pre-determined covariates are differentially missing by treatment status. Second, we turn to analyzing outcome variables.

Tables B.1 to B.6 in Appendix B present the analysis of missing covariates. The first column in each table presents the fraction missing each pre-determined control variable in the control group, while the second column presents the difference in the fraction missing between treated and control units. As indicated by the balance tests, this analysis confirms that there is no evidence of covariates being differentially missing by treatment status for the 1977 cohort, while treated units in the 1982 cohorts are around one percentage point less likely to have information on the father. Further analysis reveals that this is because control units are more likely to be in single mother households. These tests also confirm that there are no differences with respect to ETF school assignment – so randomization seems to have been successful – and the small differences we see are due to higher survey non-response by children living in households with single mothers with a lower disposable income.

Tables C.1 to C.10 in Appendix C present the analysis of missing outcomes. First, we test whether treatment status predicts whether the outcome variable is missing. It is reassuring that this does not seem to be the case. Second, we test whether observing all covariates predicts the level of each outcome variable. It appears that missing any of the control variables is related to worse outcome overall; i.e. lower grades, lower educational attainment, lower income, and a higher unemployment probability. This means that we may have unbiased (or slightly attenuated) estimates without control variables, but should be careful when interpreting the estimates including control variables – particularly those related to civil status and paternal background.

4 Empirical Results

This section presents the empirical results. We focus on shorter- and longer-term outcomes.

The short-term outcomes are 9th grade GPA, individual grades in 9th grade Math, English, and Swedish, and for the 77 cohort we also observe indicators for whether advanced Math and English were chosen in 9th grade.[18]

The long-term outcomes include four indicators for highest completed educational attainment at age 31 being at least (vocational and academic) high school and (short and long) college, respectively, as well as average yearly gross income and days unemployed during the five-year period when the individuals are 27-31 years old. Tables E.1 to E.6 in Appendix E display the empirical results. Each table presents estimates of the "survey effect" on a short- and a long-term index in order to assess whether the survey had *any* effect on educational choices and subsequent outcomes. These indices are constructed by standardizing each variable that enters the index to the control group mean and standard deviation, and then taking the average of the standardized variables with signs such that "larger is better".[19]

The total "survey effect" may be zero because of opposing effects for subgroups. We explore heterogeneous effects by differences in parental education. The salience presumably increases the most for parents who never proceeded through the educational system past compulsory schooling – around 10% of parents. However, the (dynamic) complementarity in the education production function – because of increased awareness and parental investments – may be stronger for those with more highly educated parents (Cunha and Heckman, 2007; Cunha et al., 2010; Fryer Jr et al., 2015).

We focus on the empirical strategies that provide the best overall balance; see Table 4. That is, the within-municipality specification (1) for the Swedish born sample for

[18]The short-term outcomes differ across cohorts because of institutional changes in schools and data availability. Primary through upper-secondary schooling in Sweden is regulated by the Education Act of 1985. The complete law text, and its changes over time, is available in Riksdagens law archive (link). Bjorklund et al. (2005) also provide more details on these institutional changes that also drive the increased number of schools for the 1982 cohort.

[19]This way of constructing outcome indices builds on Kling et al. (2007).

the 1977 cohort and the between-school specification (2) for the pooled sample. For comparison, however, we show both specifications for both samples. We conclude that there is overall no large effects of the survey on short- and longer-term outcomes. However, the preferred specification for the 1977 cohort (within-municipality, Tables E.1 and E.2) reveals a negative effect on grades for those whose parents have a low level of education. Responding to the survey implied 0.05 standard deviations lower short-term index, which is increasing the grade inequality slightly from a baseline of 0.43 standard deviations; see Table D.1. This negative effect on grades may even have longer-term negative effects as Table C.3 suggests that being at a surveyed school meant a 0.06 standard deviations lower long-term index as those with low parental education become 3 percentage points less likely to attain a high school degree and a short college degree. Consequently, their income is SEK 7,971-9,583 lower. This is again a small effect (less than 5%) as average income is SEK 186,203 for this group, which is SEK 31,125 lower than the average income for those with highly educated parents. For the pooled sample, however, we do not see this negative survey effect for those with low parental education, but if anything inequality is still increased as there is some indication of a positive (English) grade effect (between-school, Table E.6) and maybe even a small positive longer-term survey effect on income (within-municipality, Table E.5) for those with high parental education.

4.1 Heterogeneity by Municipality

The within-municipality estimates of δ_1 in (1) and between-school estimates of α_1 in (2) are weighted averages of the corresponding survey effect estimates within each municipality. Thus, they could mask a lot of heterogeneity. Figures F.1 and F.2 display estimates for each municipality. We split the sample of municipalities by whether all schools within the municipality had sampled classes or not. Estimated survey effects are presented by the fraction of surveyed students in the municipality. First, comparing (a) to (c) and (b) to (d), we see that the within-municipality estimates are much noisier in municipalities in which all schools are surveyed. This may be because potential spillovers from students in the same school are more prevalent in these municipalities in which

all schools have some surveyed classes. Spillovers could operate through students who share information with their friends or through teachers. Second, comparing (c) to (e) and (d) to (f) we see that the between-school estimates are attenuated towards zero compared to the within-municipality estimates in most municipalities. However, a few small municipalities exhibit positive (negative) spillovers as the between-school estimates are statistically significant, while the within-municipality estimates are not. Third, the three largest municipalities have the lowest fraction of students sampled and almost always have a precisely estimates zero effect. One exception is the significantly negative short-term effect in for the 1977 cohort; see Figure F.1 (c). Overall, there is no indication that there are systematically different treatment effects at scale.

4.2 Sibling Spillovers

We further analyze heterogeneity in survey effects and potential spillovers on non-surveyed siblings in order to better understand the mechanisms.

We might find a zero (or small) effect of being surveyed because the parental survey was assigned in 6th grade. This may be too long before students start to make consequential education choices; e.g. high school enrollment and track choice, college enrollment and major choice. To test whether this may be the case, we focus on whether the parent survey had an impact on the education choices of non-surveyed siblings. We present separate estimates for non-surveyed older siblings who were attending compulsory schooling grades $g = 7, 8, 9$ (14-16 years old) when their parent was administered a survey (at random) because they had a sibling in one of the ETF cohorts. We also present separate estimates for non-surveyed siblings who were stipulated to attend grades $g = 10, ..., 16$ (17-23 years old) and thus high school and college bound when the parent was administered the survey.[20] These siblings – especially those who are in the first three years after compulsory schooling, $g = 10, 11, 12$ – are at critical junctions on their schooling trajectory and may be particularly sensitive to the influence of their parents becoming more aware of the importance of schooling.

[20]Figure D.2 shows the distribution of older siblings by stipulated grade.

Figures G.9 to G.13 summarize the short- and long-term effects of the parent survey (on y-axis) by child grade (on x-axis) for siblings of the focal child in the sampled cohorts. These are constructed by separately estimating (1) and (2), respectively, by sibling grade. The outcome variables are those of the non-surveyed siblings, while the timing of "treatment" refers to when the parent was surveyed and the focal child was in 6th grade. First, note that the fact that there is no effect on the short-term index and grades at the end of compulsory schooling serves as a "placebo" check for siblings who are already in grade 10 and beyond. Second, there tends to be a long-term effect on siblings who completed compulsory schooling the year prior to their parent being surveyed (i.e. $g = 10$). However, the grade-by-grade estimates are generally not precise enough to draw strong conclusions. Therefore, we separately test whether siblings are particularly sensitive to education becoming more salient to their parents when they are in the first three years after compulsory schooling and stipulated to make their first critical career specialization choices. Tables G.1 to G.4 show that there is some evidence of positive spillovers on non-surveyed siblings. In families with high parental education, siblings tend to shift from academic to vocational high school track if their parent was surveyed when they were high school bound. There is some indication that the vocational track is a better "match" for these students as they have fewer unemployment days when they are 27-31 years old.

5 Conclusion

Most empirical advances in the social and health sciences over the past decades depended crucially on the use of survey data. If surveying individuals draws their attention to risks, returns, or choices previously not salient to them and this changes their subsequent behavior, then it may bias parameter estimates and conclusions drawn from survey data.

We assess whether surveys causally changed education choices and outcomes of students attending compulsory school in Sweden in the 80s and early 90s. We do not find

strong reasons to worry about extensive surveying changing education choices and subsequent labor market outcomes. This is reassuring for both the external and internal validity of estimates based on (this) survey data.

There are, however, some cases for which the survey changed grades, educational attainment, income, and employment of the surveyed individuals (and their schoolmates and siblings) in the early career. This implies that we should be cautious when administering extensive education surveys, since the surveys themselves may change the very behavior they are designed to study. Future research should aim to quantify the potential biases this change in behavior may entail for parameter estimates obtained from (this) survey data. Our results also suggest that the timing and delivery of information are important, and "nudging" in an education setting may increase inequality.

References

Adair, J. G. (1984). The hawthorne effect: a reconsideration of the methodological artifact. *Journal of applied psychology 69*(2), 334.

Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics*, 48–83.

Angrist, J. D. and V. Lavy (1999). Using maimonides'rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics 114*(2).

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*(1), 343–375.

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.

Arcidiacono, P., V. J. Hotz, and S. Kang (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics 166*(1), 3–16.

Avvisati, F., M. Gurgand, N. Guyon, and E. Maurin (2014). Getting parents involved: A field experiment in deprived schools. *The Review of Economic Studies 81*(1), 57–83.

Azmat, G., . I. N. (2010). The importance of relative performance feedback information: Evidence from a natural field experiment using high school students. *Journal of Public Economics 94*, 435–452.

Azmat, G., M. Bagues, A. Cabrales, and N. Iriberri (2016). What you don't know... can't hurt you? a field experiment on relative performance feedback in higher education.

Bailar, B. A. (1975). The effects of rotation group bias on estimates from panel surveys. *Journal of the American Statistical Association 70*(349), 23–30.

Bandiera, O., L. V. . R. I. (2017). Blissful ignorance? a natural experiment on the effect of feedback on students' performance. *Labour Economics 34*, 13–25.

Barberis, N. and R. Thaler (2003). A survey of behavioral finance. *Handbook of the Economics of Finance 1*, 1053–1128.

Bjorklund, A., M. Clark, P.-A. Edin, P. Fredriksson, and A. Kreuger (2005). The market comes to education in sweden. *Russell Sage Foundation, New York*.

Bobba, M. and V. Frisancho (2016). Learning about oneself: The effects of signalling academic ability on school choice. Technical report, Working paper.

Bound, J., C. Brown, and N. Mathiowetz (2001). Measurement error in survey data. *Handbook of econometrics 5*, 3705–3843.

Bursztyn, L., . J. R. (2017). How does peer pressure affect educational investments? *Quarterly Journal of Economics 130*(3), 1329–1367.

Carrell, S., . S. B. (2017). Why do college going interventions work. *Applied Economics 9*(3), 124–151.

Cunha, F. and J. Heckman (2007). The technology of skill formation. *The 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.

Czibor, E., S. Onderstal, R. Sloof, and M. Van Praag (2014). Does relative grading help male students? evidence from a field experiment in the classroom.

Damgaard, M. T. and H. S. Nielsen (2018). Nudging in education. *Economics of Education Review 64*, 313–342.

Davis, D., I. Jivet, R. F. Kizilcec, G. Chen, C. Hauff, and G.-J. Houben (2017). Follow the successful crowd: Raising mooc completion rates through social comparison at scale. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, LAK '17, New York, NY, USA, pp. 454–463. ACM.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature 47*(2), 315–372.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review 109*(8), 2728–2765.

Dustan, A. (2018). Family networks and school choice. *Journal of Development Economics 134*, 372–391.

Emanuelsson, I. (1979). Utvärdering genom uppföljning av elever ett nytt individualstatistikprojekt.

Facchinello, L. (2015). The impact of early grading on academic choices: Mechanisms and social implications. Job market paper, Stockholm School of Economics.

Fischer, M. and V. Wagner (2017). Effects of timing and reference frame of feedback: Evidence from a field experiment in secondary schools. Technical report, Working Paper.

Franke, R. H. and J. D. Kaul (1978). The hawthorne experiments: First statistical interpretation. *American sociological review*, 623–643.

Fredriksson, P., B. Öckert, and H. Oosterbeek (2013). Long-term effects of class size. *Quarterly Journal of Economics 128*(1), 249–285.

Fredriksson, P., B. Öckert, and H. Oosterbeek (2016). Parental responses to public investments in children: Evidence from a maximum class size rule. *Journal of Human Resources 51*(4), 832–868.

Fryer Jr, R. G., S. D. Levitt, and J. A. List (2015). Parental incentives and early childhood achievement: A field experiment in chicago heights. Technical report, National Bureau of Economic Research.

Goux, D., M. Gurgand, and E. Maurin (2014). Adjusting your dreams? highschool plans and dropout behavior. *IZA discussion paper No. 7948*.

Harackiewicz, J. M., R. C. S., C. S. Hulleman, and J. S. Hyde (2008). Helping parents to motivate adolescents in mathematics and science: An experimental test of a utility-value intervention. *Psychological Science 23*(8), 899–906.

Härnqvist, K. (1998). A longitudinal program for studying education and career development.

Hastings, J., C. A. Neilson, and S. D. Zimmerman (2015). The effects of earnings disclosure on college enrollment decisions. Working paper, National Bureau of Economic Research.

Hastings, J. S. and J. M. Weinstein (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly journal of economics 123*(4), 1373–1414.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

Jalava, N., J. S. Joensen, and E. Pellas (2015). Grades and rank: Impacts of non-financial incentives on test performance. *Journal of Economic Behavior & Organization 115*, 161–196.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics 125*(2), 515–548.

Joensen, J. S. and H. S. Nielsen (2018). Spillovers in education choice. *Journal of Public Economics 157*, 158 – 183.

Jones, S. R. (1992). Was there a hawthorne effect? *American Journal of sociology 98*(3), 451–468.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Koch, A., J. Nafziger, and H. S. Nielsen (2015). Behavioral economics of education. *Journal of Economic Behavior & Organization 115*, 3–17.

Lavecchia, A. M., H. Liu, and P. Oreopoulos (2016). Behavioral economics of education: Progress and possibilities. In *Handbook of the Economics of Education*, Volume 5, pp. 1–74. Elsevier.

Martinez, I. (2014). The effects of nudges on students' effort and performance: Lessons from a mooc. Working paper, EdPolicyWorks.

McCarney, R., J. Warner, S. Iliffe, R. Van Haselen, M. Griffin, and P. Fisher (2007). The hawthorne effect: a randomised, controlled trial. *BMC medical research methodology 7*(1), 30.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from madagascar. *MIT mimeo*.

Parsons, H. M. (1974). What happened at hawthorne?: New evidence suggests the hawthorne effect resulted from operant reinforcement contingencies. *Science 183*(4128), 922–932.

Pekkala Kerr, S., T. Pekkarinen, R. Uusitalo, et al. (2015). Post-secondary education and information on labor market prospects: A randomized field experiment.

Pistolesi, N. (2017). The effect of advising students at college entrance: evidence from a french university reform. *Labour Economics 44*, 106–121.

Qureshi, J. A. (2018). Additional returns to investing in girls' education: Impact on younger sibling human capital. *The Economic Journal 128*(616), 3285–3319.

Roethlisberger, F. J. and W. J. Dickson (1939). *Management and the Worker*, Volume 5. Psychology press.

Simon, H. A. (1955). A behavioral model of rational choice. *the Quarterly Journal of Economics*, 99–118.

Solon, G. (1986). Effects of rotation group bias on estimation of unemployment. *Journal of Business & Economic Statistics 4*(1), 105–109.

Tran, A., . Z. R. (2012). Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics 96*, 645–650.

Wagner, V., . R. G. (2015). Peers or parents? on non-monetary incentives in schools. Working Paper 203, DICE.

Wiswall, M. and B. Zafar (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies 82*(2), 791–824.

Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics 29*(2), 301–348.

Zwane, A. P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D. S. Karlan, R. Hornbeck, X. Giné, E. Duflo, F. Devoto, B. Crepon, and A. Banerjee (2011). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences 108*(5), 1821–1826.

Appendices

Appendices are organized as follows: Appendix A presents balance tests for each cohort and empirical strategy, Appendix B shows evidence on covariates missing at random, Appendix C relates missing variables to outcomes, Appendix D presents additional descriptives on sample characteristics and outcomes variables, Appendix E presents the main empirical estimates of the survey effect on the focal child, Appendix F presents estimates by municipality, while Appendix G presents the empirical estimates of the spillover effects of the parent survey on non-surveyed children; i.e. the siblings of the focal child.

A Balance Tests

Table A.1:
Differences in student background by treatment status:
Within municipality specification - 1977 cohort

	Control mean	T-C difference All	Restricted	Hedge's g
Female	0.490	0.005	0.015	-0.001
	(0.500)	(0.010)	(0.011)	
Number of siblings	1.995	-0.000	-0.009	-0.022
	(1.451)	(0.029)	(0.029)	
Foreign born	0.053	0.003	0.000	0.072
	(0.225)	(0.004)		
Swedish born, foreign parent	0.112	0.001	0.000	0.085
	(0.315)	(0.006)		
Divorced parent	0.188	-0.016*	-0.012	0.134
	(0.391)	(0.008)	(0.008)	
Father age at birth	30.237	0.044	0.011	0.051
	(5.690)	(0.132)	(0.154)	
Father disposable income	180.309	-2.389	-2.821	0.066
	(84.845)	(1.810)	(1.798)	
Father educ: high school	0.429	0.000	0.006	-0.022
	(0.495)	(0.010)	(0.012)	
Father educ: college	0.289	-0.007	-0.007	0.096
	(0.453)	(0.012)	(0.013)	
Father: in the labor force	0.947	0.002	0.002	-0.066
	(0.223)	(0.004)	(0.004)	
Mother age at birth	27.445	0.085	0.063	0.034
	(4.951)	(0.123)	(0.133)	
Mother disposable income	130.277	0.405	0.420	0.138
	(51.105)	(1.084)	(1.202)	
Mother educ: high school	0.436	-0.001	0.003	-0.057
	(0.496)	(0.011)	(0.012)	
Mother educ: college	0.307	-0.007	-0.004	0.095
	(0.461)	(0.012)	(0.013)	
Mother: in the labor force	0.939	0.001	0.006	-0.042
	(0.240)	(0.004)	(0.004)	
Hotelling's T-squared		154.14	120.36	
Observations		18162	14322	
Schools		233	230	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between treated and control units, within municipality. All regressions control for school size. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table A.2:
Differences in mother and child health by treatment status:
Within municipality specification - 1977 cohort

	Control mean	T-C difference All	Restricted	Hedge's g
Healthy mother, no diagnosis	0.637	-0.001	0.001	-0.003
	(0.481)	(0.010)	(0.010)	
Length of pregnancy (weeks)	39.696	0.011	0.040	-0.007
	(1.793)	(0.033)	(0.035)	
Healthy child, no diagnosis	0.660	0.010	0.010	-0.088
	(0.474)	(0.009)	(0.009)	
APGAR 1 min (0-10)	8.796	-0.000	-0.006	-0.040
	(1.223)	(0.025)	(0.028)	
Birth weight (g)	3,480.030	-2.545	-0.540	-0.016
	(535.616)	(10.722)	(11.107)	
Low birth weight (≤2500g)	0.037	-0.002	-0.002	0.015
	(0.189)	(0.003)	(0.004)	
Very low birth weight (≤1500g)	0.003	0.001	0.001	-0.024
	(0.050)	(0.001)	(0.001)	
Extremely low birth weight (≤1000g)	0.000	-0.000	-0.000	-0.001
	(0.016)	(0.000)	(0.000)	
Small for Gestational Age (SGA)	0.040	0.002	0.002	0.000
	(0.197)	(0.003)	(0.004)	
Large for Gestational Age (LGA)	0.023	0.000	0.002	-0.017
	(0.149)	(0.003)	(0.003)	
Hotelling's T-squared		27.19	23.50	
Observations		19055	15410	
Schools		233	231	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between treated and control units, within municipality. All regressions control for school size. Data is available only for Swedish born. Column 2 considers the full sample. Column 3 restricts to children with Swedish parents. Standard errors clustered at the school level.

Table A.3:
Differences in student background by treatment assignment:
Between school specification - 1977 cohort

	NS mean	S - NS difference	Hedge's g
Female	0.496	-0.005	0.022
	(0.500)	(0.010)	
Number of siblings	1.970	0.084	-0.038
	(1.458)	(0.051)	
Foreign born	0.060	0.017*	0.075
	(0.237)	(0.009)	
Swedish born, foreign parent	0.131	-0.013	0.141
	(0.338)	(0.009)	
Divorced parent	0.201	0.026*	0.105
	(0.401)	(0.014)	
Father age at birth	30.467	0.053	0.088
	(5.768)	(0.236)	
Father disposable income	183.017	-3.511	0.081
	(97.710)	(5.214)	
Father educ: high school	0.424	-0.014	-0.026
	(0.494)	(0.017)	
Father educ: college	0.316	0.001	0.137
	(0.465)	(0.030)	
Father: in the labor force	0.941	-0.010	-0.076
	(0.236)	(0.007)	
Mother age at birth	27.661	-0.138	0.088
	(5.002)	(0.241)	
Mother disposable income	134.318	2.650	0.187
	(52.064)	(1.745)	
Mother educ: high school	0.420	-0.006	-0.076
	(0.494)	(0.017)	
Mother educ: college	0.337	-0.014	0.145
	(0.473)	(0.030)	
Mother: in the labor force	0.935	-0.007	-0.043
	(0.247)	(0.008)	
Hotelling's T-squared		438.75	
Observations		18162	
Schools		233	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between students graduating from sampled and non-sampled schools, within municipality. All regressions control for school size. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table A.4:
Differences in mother and child health by treatment assignment:
Between school specification - 1977 cohort

	NS mean	S - NS difference	Hedge's g
Healthy mother, no diagnosis	0.634	-0.003	-0.011
	(0.482)	(0.009)	
Length of pregnancy (weeks)	39.677	0.015	-0.021
	(1.789)	(0.041)	
Healthy child, no diagnosis	0.648	-0.013	-0.074
	(0.478)	(0.012)	
APGAR 1 min (0-10)	8.749	0.018	-0.079
	(1.260)	(0.034)	
Birth weight (g)	3,470.439	3.326	-0.037
	(533.638)	(13.124)	
Low birth weight ($\leq$2500g)	0.037	0.002	0.009
	(0.190)	(0.004)	
Very low birth weight ($\leq$1500g)	0.002	-0.000	-0.012
	(0.049)	(0.001)	
Extremely low birth weight ($\leq$1000g)	0.000	0.000	-0.015
	(0.011)	(0.000)	
Small for Gestational Age (SGA)	0.042	0.001	0.016
	(0.201)	(0.004)	
Large for Gestational Age (LGA)	0.022	0.000	-0.015
	(0.146)	(0.003)	
Hotelling's T-squared		56.84	
Observations		19055	
Schools		233	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between students graduating from sampled and non-sampled schools, within municipality. All regressions control for school size. Data is only available for Swedish born. Standard errors clustered at the school level.

Table A.5:
Differences in student background by treatment status:
Within municipality specification - 1982 cohort

	Control mean	T-C difference All	Restrict 1	Restrict 2	Hedge's g
Female	0.497	-0.001	-0.003	0.002	0.006
	(0.500)	(0.008)	(0.009)	(0.008)	
Number of siblings	2.119	-0.077***	-0.056**	-0.051**	0.015
	(1.490)	(0.025)	(0.025)	(0.026)	
Foreign born	0.082	-0.005	0.000	-0.003	0.072
	(0.274)	(0.004)		(0.004)	
Swedish born, foreign parent	0.112	0.007	0.000	0.008	0.012
	(0.316)	(0.005)		(0.005)	
Divorced parent	0.127	-0.013***	-0.006	-0.008*	0.074
	(0.333)	(0.005)	(0.006)	(0.005)	
Father age at birth	31.452	0.059	0.083	0.020	0.053
	(5.763)	(0.099)	(0.109)	(0.102)	
Father disposable income	207.583	5.488	4.965	5.850	0.016
	(145.449)	(3.718)	(4.528)	(3.938)	
Father educ: high school	0.409	0.006	0.009	0.008	-0.064
	(0.492)	(0.008)	(0.010)	(0.009)	
Father educ: college	0.362	0.008	-0.005	0.006	0.098
	(0.481)	(0.011)	(0.012)	(0.011)	
Father: in the labor force	0.939	0.007*	0.004	0.005	-0.063
	(0.239)	(0.004)	(0.003)	(0.004)	
Mother age at birth	28.608	0.133	0.100	0.086	0.053
	(5.151)	(0.089)	(0.097)	(0.095)	
Mother disposable income	163.271	-1.791*	-2.610**	-1.340	0.102
	(73.342)	(1.032)	(1.102)	(1.059)	
Mother educ: high school	0.421	0.012	0.013	0.011	-0.088
	(0.494)	(0.009)	(0.010)	(0.009)	
Mother educ: college	0.362	0.010	-0.001	0.007	0.085
	(0.481)	(0.010)	(0.011)	(0.011)	
Mother: in the labor force	0.912	0.008	0.002	0.006	-0.032
	(0.283)	(0.005)	(0.004)	(0.005)	
Hotelling's T-squared		157.83	154.21	133.70	
Observations		23191	17712	21505	
Schools		348	339	343	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between treated and control units, within municipality. All regressions control for school size. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table A.6:
Differences in mother and child health by treatment status:
Within municipality specification - 1982 cohort

	Control mean	T-C difference All	Restrict 1	Restrict 2	Hedge's g
Healthy mother, no diagnosis	0.568	0.002	0.002	0.003	-0.020
	(0.495)	(0.008)	(0.009)	(0.009)	
Length of pregnancy (weeks)	39.448	0.054*	0.049	0.045	-0.027
	(1.830)	(0.028)	(0.031)	(0.031)	
Healthy child, no diagnosis	0.689	-0.001	-0.000	-0.004	-0.019
	(0.463)	(0.007)	(0.008)	(0.007)	
APGAR 1 min (0-10)	8.798	0.019	0.011	0.023	-0.043
	(1.117)	(0.018)	(0.020)	(0.019)	
Birth weight (g)	3,486.083	15.515*	14.579	12.724	-0.038
	(549.400)	(8.859)	(9.660)	(9.040)	
Low birth weight (≤2500g)	0.040	-0.003	-0.003	-0.002	0.012
	(0.195)	(0.003)	(0.003)	(0.003)	
Very low birth weight (≤1500g)	0.004	-0.001	-0.000	-0.001	0.014
	(0.066)	(0.001)	(0.001)	(0.001)	
Extremely low birth weight (≤1000g)	0.001	-0.000	-0.000	-0.000	0.009
	(0.031)	(0.000)	(0.000)	(0.000)	
Small for Gestational Age (SGA)	0.034	-0.005**	-0.005*	-0.005*	0.031
	(0.180)	(0.002)	(0.003)	(0.003)	
Large for Gestational Age (LGA)	0.027	-0.004	-0.003	-0.005*	0.000
	(0.161)	(0.003)	(0.003)	(0.003)	
Hotelling's T-squared		17.58	11.36	17.74	
Observations		22064	17712	19225	
Schools		349	341	345	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between treated and control units, within municipality. All regressions control for school size. Data is available only for Swedish born. Column 2 considers the full sample. Column 3 restricts to children with Swedish parents. Standard errors clustered at the school level.

Table A.7:
Differences in student background by treatment assignment:
Between school specification - 1982 cohort

	NS mean	S - NS difference	Hedge's g
Female	0.494	0.005	-0.005
	(0.500)	(0.010)	
Number of siblings	2.148	-0.089**	0.033
	(1.527)	(0.036)	
Foreign born	0.095	-0.006	0.099
	(0.294)	(0.007)	
Swedish born, foreign parent	0.119	0.007	0.032
	(0.323)	(0.007)	
Divorced parent	0.139	-0.007	0.079
	(0.346)	(0.008)	
Father age at birth	31.578	0.007	0.051
	(5.904)	(0.149)	
Father disposable income	209.580	1.387	0.023
	(218.803)	(6.052)	
Father educ: high school	0.404	0.002	-0.039
	(0.491)	(0.013)	
Father educ: college	0.366	0.013	0.049
	(0.482)	(0.020)	
Father: in the labor force	0.934	0.002	-0.057
	(0.249)	(0.005)	
Mother age at birth	28.619	0.141	0.024
	(5.251)	(0.165)	
Mother disposable income	166.038	-0.369	0.095
	(68.060)	(1.609)	
Mother educ: high school	0.404	0.015	-0.081
	(0.491)	(0.012)	
Mother educ: college	0.369	0.011	0.055
	(0.483)	(0.020)	
Mother: in the labor force	0.903	0.011	-0.057
	(0.296)	(0.008)	
Hotelling's T-squared		127.66	
Observations		23191	
Schools		348	

* p < 0.10, ** p < 0.05, *** p < 0.01

The table tests for pre-treatment differences between students graduating from sampled and non-sampled schools, within municipality. All regressions control for school size. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table A.8:
Differences in mother and child health by treatment assignment:
Between school specification - 1982 cohort

	NS mean	S - NS difference	Hedge's g
Healthy mother, no diagnosis	0.568	-0.007	-0.009
	(0.495)	(0.008)	
Length of pregnancy (weeks)	39.440	0.067**	-0.017
	(1.831)	(0.031)	
Healthy child, no diagnosis	0.695	-0.007	0.010
	(0.460)	(0.008)	
APGAR 1 min (0-10)	8.771	-0.004	-0.050
	(1.165)	(0.018)	
Birth weight (g)	3,480.649	5.361	-0.029
	(552.534)	(9.615)	
Low birth weight (≤2500g)	0.042	-0.005*	0.020
	(0.200)	(0.003)	
Very low birth weight (≤1500g)	0.003	-0.000	-0.015
	(0.058)	(0.001)	
Extremely low birth weight (≤1000g)	0.000	0.000	-0.018
	(0.022)	(0.000)	
Small for Gestational Age (SGA)	0.033	0.001	0.011
	(0.180)	(0.003)	
Large for Gestational Age (LGA)	0.029	-0.005*	0.018
	(0.167)	(0.003)	
Hotelling's T-squared		15.38	
Observations		22064	
Schools		349	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between students graduating from sampled and non-sampled schools, within municipality. All regressions control for school size. Data is only available for Swedish born. Standard errors clustered at the school level.

Table A.9:
Differences in student background by treatment status:
Within municipality specification - All

	Control mean	T-C difference All	Restricted	Hedge's g
Female	0.494	0.002	0.004	0.002
	(0.500)	(0.006)	(0.007)	
Number of siblings	2.060	-0.045**	-0.038**	-0.010
	(1.473)	(0.020)	(0.019)	
Foreign born	0.068	-0.002	0.000	0.057
	(0.252)	(0.003)		
Swedish born, foreign parent	0.112	0.005	0.000	0.036
	(0.315)	(0.004)		
Divorced parent	0.156	-0.014***	-0.008*	0.119
	(0.363)	(0.004)	(0.005)	
Father age at birth	30.873	0.042	0.047	0.021
	(5.760)	(0.078)	(0.090)	
Father disposable income	194.606	2.485	1.890	-0.004
	(121.243)	(2.558)	(3.044)	
Father educ: high school	0.419	0.004	0.009	-0.044
	(0.493)	(0.007)	(0.008)	
Father educ: college	0.327	0.002	-0.006	0.075
	(0.469)	(0.008)	(0.009)	
Father: in the labor force	0.943	0.005*	0.003	-0.059
	(0.232)	(0.003)	(0.002)	
Mother age at birth	28.054	0.102	0.075	0.014
	(5.090)	(0.072)	(0.080)	
Mother disposable income	147.562	-1.030	-1.501*	0.035
	(65.825)	(0.782)	(0.838)	
Mother educ: high school	0.428	0.007	0.008	-0.073
	(0.495)	(0.007)	(0.008)	
Mother educ: college	0.336	0.003	-0.003	0.072
	(0.472)	(0.008)	(0.009)	
Mother: in the labor force	0.925	0.005	0.003	-0.020
	(0.264)	(0.004)	(0.003)	
Hotelling's T-squared		217.15	167.98	
Observations		41353	32034	
Schools		478	469	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between treated and control units, within municipality. All regressions control for school size. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table A.10:
Differences in mother and child health by treatment status:
Within municipality specification - All

	Control mean	T-C difference All	Restricted	Hedge's g
Healthy mother, no diagnosis	0.602	0.000	0.000	0.006
	(0.490)	(0.006)	(0.007)	
Length of pregnancy (weeks)	39.570	0.037*	0.043*	0.000
	(1.816)	(0.021)	(0.023)	
Healthy child, no diagnosis	0.675	0.002	0.002	-0.052
	(0.469)	(0.006)	(0.006)	
APGAR 1 min (0-10)	8.797	0.010	0.003	-0.042
	(1.170)	(0.015)	(0.016)	
Birth weight (g)	3,483.124	7.554	7.204	-0.033
	(542.704)	(6.835)	(7.301)	
Low birth weight (≤2500g)	0.038	-0.003	-0.002	0.011
	(0.192)	(0.002)	(0.002)	
Very low birth weight (≤1500g)	0.003	0.000	0.000	-0.002
	(0.059)	(0.001)	(0.001)	
Extremely low birth weight (≤1000g)	0.001	-0.000	-0.000	0.002
	(0.025)	(0.000)	(0.000)	
Small for Gestational Age (SGA)	0.037	-0.002	-0.001	0.025
	(0.189)	(0.002)	(0.002)	
Large for Gestational Age (LGA)	0.025	-0.002	-0.001	-0.009
	(0.155)	(0.002)	(0.002)	
Hotelling's T-squared		46.19	36.55	
Observations		41119	33122	
Schools		479	472	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between treated and control units, within municipality. All regressions control for school size. Data is only available for Swedish born. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Standard errors clustered at the school level.

Table A.11:
Differences in student background by treatment assignment:
Between school specification - All

	NS mean	S - NS difference	Hedge's g
Female	0.495	0.002	0.006
	(0.500)	(0.008)	
Number of siblings	2.048	-0.027	-0.016
	(1.491)	(0.032)	
Foreign born	0.075	0.003	0.065
	(0.264)	(0.006)	
Swedish born, foreign parent	0.126	-0.001	0.080
	(0.332)	(0.006)	
Divorced parent	0.174	0.004	0.121
	(0.379)	(0.006)	
Father age at birth	30.952	0.020	0.029
	(5.854)	(0.126)	
Father disposable income	194.637	-0.385	-0.001
	(162.744)	(3.879)	
Father educ: high school	0.415	-0.005	-0.027
	(0.493)	(0.010)	
Father educ: college	0.338	0.010	0.062
	(0.473)	(0.017)	
Father: in the labor force	0.938	-0.002	-0.058
	(0.242)	(0.004)	
Mother age at birth	28.079	0.044	0.013
	(5.134)	(0.131)	
Mother disposable income	148.175	0.807	0.028
	(61.623)	(1.239)	
Mother educ: high school	0.413	0.009	-0.073
	(0.492)	(0.010)	
Mother educ: college	0.351	0.002	0.076
	(0.477)	(0.017)	
Mother: in the labor force	0.921	0.004	-0.030
	(0.270)	(0.006)	
Hotelling's T-squared		295.80	
Observations		41353	
Schools		478	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between students graduating from sampled and non-sampled schools, within municipality. All regressions control for school size. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table A.12:
Differences in mother and child health by treatment assignment:
Between school specification - All

	NS mean	S - NS difference	Hedge's g
Healthy mother, no diagnosis	0.606	-0.007	0.016
	(0.489)	(0.006)	
Length of pregnancy (weeks)	39.578	0.047*	0.007
	(1.810)	(0.025)	
Healthy child, no diagnosis	0.667	-0.007	-0.043
	(0.471)	(0.007)	
APGAR 1 min (0-10)	8.758	0.007	-0.065
	(1.221)	(0.017)	
Birth weight (g)	3,474.752	5.098	-0.035
	(541.706)	(7.387)	
Low birth weight (≤2500g)	0.039	-0.002	0.012
	(0.194)	(0.003)	
Very low birth weight (≤1500g)	0.003	-0.000	-0.017
	(0.053)	(0.001)	
Extremely low birth weight (≤1000g)	0.000	0.000	-0.020
	(0.017)	(0.000)	
Small for Gestational Age (SGA)	0.039	0.001	0.022
	(0.193)	(0.002)	
Large for Gestational Age (LGA)	0.025	-0.003	-0.003
	(0.155)	(0.002)	
Hotelling's T-squared		77.88	
Observations		41119	
Schools		479	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests for pre-treatment differences between students graduating from sampled and non-sampled schools, within municipality. All regressions control for school size. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

B Missing Covariates

Table B.1: Missingness in pre-treatment variables: Within-municipality specification - 1977 cohort

	Control mean	T-C difference All	T-C difference Restricted	Hedge's g
Female	0.000	0.000	0.000	
	(0.000)			
Number of siblings	0.000	0.000	0.000	
	(0.000)			
Foreign born	0.000	0.000	0.000	
	(0.000)			
Swedish born, foreign parent	0.000	0.000	0.000	
	(0.000)			
Divorced parent	0.075	-0.000	-0.002	0.014
	(0.263)	(0.005)	(0.006)	
Father age at birth	0.077	-0.000	-0.003	0.030
	(0.266)	(0.005)	(0.005)	
Father disposable income	0.089	-0.002	-0.004	0.041
	(0.284)	(0.006)	(0.006)	
Father educ: high school	0.116	-0.006	-0.006	0.077
	(0.320)	(0.007)	(0.006)	
Father educ: college	0.116	-0.006	-0.006	0.077
	(0.320)	(0.007)	(0.006)	
Father: in the labor force	0.104	-0.005	-0.006	0.058
	(0.305)	(0.006)	(0.006)	
Mother age at birth	0.064	-0.000	-0.002	0.013
	(0.244)	(0.005)	(0.005)	
Mother disposable income	0.067	0.000	-0.002	0.011
	(0.249)	(0.005)	(0.005)	
Mother educ: high school	0.078	0.001	-0.002	0.030
	(0.268)	(0.005)	(0.005)	
Mother educ: college	0.078	0.001	-0.002	0.030
	(0.268)	(0.005)	(0.005)	
Mother: in the labor force	0.072	-0.001	-0.002	0.021
	(0.258)	(0.005)	(0.005)	
Any tested variable	0.130	-0.006	-0.007	0.086
	(0.337)	(0.007)	(0.007)	
Hotelling's T-squared		154.14	120.36	
Observations		20759	15909	
Schools		233	231	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests whether treatment assignment predicts missingness of pre-treatment variables, within municipality. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table B.2: Missingness in pre-treatment variables: Between-school specification - 1977 cohort

	NS mean	S - NS difference	Hedge's g
Female	0.000	0.000	
	(0.000)		
Number of siblings	0.000	0.000	
	(0.000)		
Foreign born	0.000	0.000	
	(0.000)		
Swedish born, foreign parent	0.000	0.000	
	(0.000)		
Divorced parent	0.076	0.003	0.014
	(0.265)	(0.005)	
Father age at birth	0.079	0.009	0.023
	(0.269)	(0.006)	
Father disposable income	0.092	0.008	0.036
	(0.290)	(0.006)	
Father educ: high school	0.126	0.007	0.080
	(0.332)	(0.009)	
Father educ: college	0.126	0.007	0.080
	(0.332)	(0.009)	
Father: in the labor force	0.110	0.008	0.052
	(0.313)	(0.008)	
Mother age at birth	0.064	0.004	0.008
	(0.245)	(0.004)	
Mother disposable income	0.067	0.006	0.004
	(0.250)	(0.004)	
Mother educ: high school	0.082	0.008*	0.035
	(0.274)	(0.005)	
Mother educ: college	0.082	0.008*	0.035
	(0.274)	(0.005)	
Mother: in the labor force	0.073	0.006	0.015
	(0.260)	(0.004)	
Any tested variable	0.142	0.010	0.092
	(0.350)	(0.010)	
Hotelling's T-squared		438.75	
Observations		20759	
Schools		233	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests whether graduating from sampled and non-sampled schools predicts missingness of pre-treatment variables, within municipality. Column 2 considers the full sample. Column 3 restricts to municipalities where less than 85% of the students were sampled. Controls: school size. Prices adjusted to 2014: 1 SEK = 0.11 . Standard errors clustered at the school level.

Table B.3: Missingness in pre-treatment variables: Within-municipality specification - 1982 cohort

	Control mean	T-C difference All	Restrict 1	Restrict 2	Hedge's g
Female	0.000	0.000	0.000	0.000	
	(0.000)				
Number of siblings	0.000	0.000	0.000	0.000	
	(0.000)				
Foreign born	0.000	-0.000	0.000	-0.000	0.012
	(0.010)	(0.000)		(0.000)	
Swedish born, foreign parent	0.000	-0.000	0.000	-0.000	0.012
	(0.010)	(0.000)		(0.000)	
Divorced parent	0.073	-0.003	-0.005	-0.000	0.011
	(0.260)	(0.004)	(0.004)	(0.001)	
Father age at birth	0.077	-0.004	-0.006	-0.001	0.023
	(0.267)	(0.004)	(0.004)	(0.002)	
Father disposable income	0.087	-0.007*	-0.008*	-0.003*	0.034
	(0.282)	(0.004)	(0.004)	(0.002)	
Father educ: high school	0.111	-0.009**	-0.009*	-0.004	0.053
	(0.314)	(0.005)	(0.005)	(0.003)	
Father educ: college	0.111	-0.009**	-0.009*	-0.004	0.053
	(0.314)	(0.005)	(0.005)	(0.003)	
Father: in the labor force	0.096	-0.006	-0.008*	-0.002	0.035
	(0.294)	(0.004)	(0.004)	(0.002)	
Mother age at birth	0.064	-0.002	-0.005	-0.000	0.010
	(0.245)	(0.004)	(0.004)	(0.001)	
Mother disposable income	0.066	-0.001	-0.003	0.001	0.006
	(0.248)	(0.004)	(0.004)	(0.001)	
Mother educ: high school	0.087	-0.005	-0.003	-0.001	0.032
	(0.282)	(0.004)	(0.004)	(0.002)	
Mother educ: college	0.087	-0.005	-0.003	-0.001	0.032
	(0.282)	(0.004)	(0.004)	(0.002)	
Mother: in the labor force	0.072	-0.002	-0.003	0.002	0.012
	(0.259)	(0.004)	(0.004)	(0.001)	
Any tested variable	0.126	-0.011**	-0.008	-0.005	0.064
	(0.331)	(0.005)	(0.005)	(0.004)	
Hotelling's T-squared		157.83	154.21	133.70	
Observations		26340	19348	22783	
Schools		351	342	346	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests whether treatment assignment predicts missingness of pre-treatment variables, within municipality. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Column 4 restricts to students living in the same municipality in grades 3 and 6. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table B.4: Missingness in pre-treatment variables:
Between-school specification - 1982 cohort

	NS mean	S - NS difference	Hedge's g
Female	0.000	0.000	
	(0.000)		
Number of siblings	0.000	0.000	
	(0.000)		
Foreign born	0.000	-0.000	0.033
	(0.017)	(0.000)	
Swedish born, foreign parent	0.000	-0.000	0.033
	(0.017)	(0.000)	
Divorced parent	0.074	-0.002	0.012
	(0.262)	(0.004)	
Father age at birth	0.081	-0.005	0.028
	(0.273)	(0.004)	
Father disposable income	0.091	-0.006	0.035
	(0.288)	(0.004)	
Father educ: high school	0.121	-0.010*	0.063
	(0.326)	(0.006)	
Father educ: college	0.121	-0.010*	0.063
	(0.326)	(0.006)	
Father: in the labor force	0.103	-0.009*	0.047
	(0.304)	(0.005)	
Mother age at birth	0.065	-0.002	0.010
	(0.247)	(0.004)	
Mother disposable income	0.068	-0.003	0.013
	(0.252)	(0.004)	
Mother educ: high school	0.094	-0.010*	0.050
	(0.293)	(0.005)	
Mother educ: college	0.094	-0.010*	0.050
	(0.293)	(0.005)	
Mother: in the labor force	0.075	-0.004	0.019
	(0.264)	(0.004)	
Any tested variable	0.141	-0.017**	0.088
	(0.348)	(0.007)	
Hotelling's T-squared		127.66	
Observations		26340	
Schools		351	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests whether graduating from sampled and non-sampled schools predicts missingness of pre-treatment variables, within municipality. Column 2 considers the full sample. Column 3 restricts to municipalities where less than 85% of the students were sampled. Controls: school size. Prices adjusted to 2014: 1 SEK = 0.11 . Standard errors clustered at the school level.

Table B.5: Missingness in pre-treatment variables: Within-municipality specification - 1977 and 1982 cohort

	Control mean	T-C difference All	Restricted	Hedge's g
Female	0.000	0.000	0.000	
	(0.000)			
Number of siblings	0.000	0.000	0.000	
	(0.000)			
Foreign born	0.000	-0.000	0.000	0.009
	(0.008)	(0.000)		
Swedish born, foreign parent	0.000	-0.000	0.000	0.009
	(0.008)	(0.000)		
Divorced parent	0.074	-0.002	-0.004	0.013
	(0.261)	(0.003)	(0.003)	
Father age at birth	0.077	-0.003	-0.005	0.025
	(0.267)	(0.003)	(0.003)	
Father disposable income	0.088	-0.005	-0.007*	0.037
	(0.283)	(0.003)	(0.003)	
Father educ: high school	0.114	-0.008**	-0.008**	0.063
	(0.317)	(0.004)	(0.004)	
Father educ: college	0.114	-0.008**	-0.008**	0.063
	(0.317)	(0.004)	(0.004)	
Father: in the labor force	0.100	-0.006	-0.008**	0.047
	(0.300)	(0.004)	(0.004)	
Mother age at birth	0.064	-0.002	-0.004	0.011
	(0.245)	(0.003)	(0.003)	
Mother disposable income	0.066	-0.001	-0.003	0.008
	(0.249)	(0.003)	(0.003)	
Mother educ: high school	0.083	-0.003	-0.003	0.026
	(0.275)	(0.003)	(0.003)	
Mother educ: college	0.083	-0.003	-0.003	0.026
	(0.275)	(0.003)	(0.003)	
Mother: in the labor force	0.072	-0.001	-0.003	0.015
	(0.259)	(0.003)	(0.003)	
Any tested variable	0.128	-0.009**	-0.008**	0.073
	(0.334)	(0.004)	(0.004)	
Hotelling's T-squared		217.15	167.98	
Observations		47099	35257	
Schools		481	473	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests whether treatment assignment predicts missingness of pre-treatment variables, within municipality. Column 2 considers the full sample. Column 3 restricts to Swedish born with Swedish parents. Prices adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table B.6: Missingness in pre-treatment variables:
Between-school specification - 1977 and 1982 cohort

	NS mean	S - NS difference	Hedge's g
Female	0.000	0.000	
	(0.000)		
Number of siblings	0.000	0.000	
	(0.000)		
Foreign born	0.000	-0.000	0.019
	(0.011)	(0.000)	
Swedish born, foreign parent	0.000	-0.000	0.019
	(0.011)	(0.000)	
Divorced parent	0.075	-0.000	0.014
	(0.264)	(0.003)	
Father age at birth	0.080	-0.000	0.024
	(0.271)	(0.003)	
Father disposable income	0.092	-0.001	0.036
	(0.289)	(0.003)	
Father educ: high school	0.124	-0.003	0.072
	(0.329)	(0.005)	
Father educ: college	0.124	-0.003	0.072
	(0.329)	(0.005)	
Father: in the labor force	0.107	-0.003	0.053
	(0.309)	(0.004)	
Mother age at birth	0.065	-0.000	0.009
	(0.246)	(0.003)	
Mother disposable income	0.067	0.000	0.009
	(0.250)	(0.003)	
Mother educ: high school	0.087	-0.003	0.036
	(0.282)	(0.004)	
Mother educ: college	0.087	-0.003	0.036
	(0.282)	(0.004)	
Mother: in the labor force	0.074	-0.000	0.016
	(0.262)	(0.003)	
Any tested variable	0.142	-0.007	0.091
	(0.349)	(0.006)	
Hotelling's T-squared		295.80	
Observations		47099	
Schools		481	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table tests whether graduating from sampled and non-sampled schools predicts missingness of pre-treatment variables, within municipality. Column 2 considers the full sample. Column 3 restricts to municipalities where less than 85% of the students were sampled. Controls: school size. Prices adjusted to 2014: 1 SEK = 0.11 . Standard errors clustered at the school level.

C Missing Outcomes

Table C.1: Missingness in outcomes and covariates:
Within-municipality specification - 1977 cohort

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	0.01**	-0.10***	0.00	-0.06	0.01*	-0.08**
	(0.00)	(0.02)	(0.00)	(0.04)	(0.01)	(0.03)
Advanced Math in grade 9	-0.00	-0.07***	0.00	-0.06**	-0.00	-0.06***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Advanced Eng in grade 9	-0.01	-0.05***	-0.01	0.01	-0.00	-0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Swedish grade (grade 9, 0-100)	0.01*	-2.08***	0.00	-1.08	0.01*	-2.12***
	(0.00)	(0.38)	(0.00)	(0.92)	(0.01)	(0.73)
Math grade (grade 9, 0-100)	-0.00	-2.53***	-0.00	-3.21***	-0.00	-1.54**
	(0.00)	(0.44)	(0.00)	(1.01)	(0.00)	(0.71)
English grade (grade 9, 0-100)	0.00	-1.30***	0.00	1.11	0.00	-0.88
	(0.00)	(0.39)	(0.00)	(0.93)	(0.00)	(0.78)
GPA (grade 9, 0-100)	0.00	-2.98***	0.00	-2.87***	0.00	-3.22***
		(0.38)		(0.89)		(0.63)
Long-term index (std munic)	0.01**	-0.15***	0.00	-0.18***	0.01*	-0.15***
	(0.00)	(0.02)	(0.00)	(0.03)	(0.00)	(0.03)
Attains high school (age 31)	0.00	-0.05***	0.01	-0.04***	0.00	-0.06***
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Attains academic HS (age 31)	0.01	-0.06***	0.00	-0.02	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Attains short college (age 31)	0.00	-0.08***	0.01	-0.08***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.02)
Attains college (age 31)	0.00	-0.07***	0.01	-0.07***	0.00	-0.00
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Gross income (27-31 yrs old)	0.01*	-217.76***	0.00	-400.80***	0.01	-229.61***
	(0.00)	(29.87)	(0.00)	(61.78)	(0.00)	(49.40)
Days/year unemp. (27-31 yrs old)	0.01*	2.86***	0.00	2.32	0.01	6.34***
	(0.00)	(0.78)	(0.00)	(1.85)	(0.00)	(1.80)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table C.2: Missingness in outcomes and covariates:
Within-municipality specification - 1977 cohort - Swedish born

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	0.00	-0.07***	0.00	-0.03	0.00	-0.10**
	(0.00)	(0.02)	(0.00)	(0.06)	(0.00)	(0.04)
Advanced Math in grade 9	-0.00	-0.05***	-0.01	-0.05	-0.00	-0.06**
	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)	(0.03)
Advanced Eng in grade 9	-0.01	-0.04***	-0.01	-0.01	-0.00	-0.03
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)
Swedish grade (grade 9, 0-100)	-0.00	-1.35***	0.00	0.01	-0.00	-2.48**
	(0.00)	(0.48)	(0.00)	(1.25)	(0.00)	(0.96)
Math grade (grade 9, 0-100)	-0.00	-1.71***	0.00	-1.66	-0.00	-1.15
	(0.00)	(0.53)	(0.00)	(1.43)	(0.00)	(1.05)
English grade (grade 9, 0-100)	0.00	-0.70	0.00	0.10	0.00	-1.18
	(0.00)	(0.45)	(0.00)	(1.27)	(0.00)	(1.01)
GPA (grade 9, 0-100)	0.00	-1.76***	0.00	-2.23	0.00	-3.06***
		(0.42)		(1.40)		(0.80)
Long-term index (std munic)	0.00	-0.09***	0.00	-0.08*	0.00	-0.13***
	(0.00)	(0.02)	(0.00)	(0.05)	(0.00)	(0.04)
Attains high school (age 31)	0.00	-0.03***	0.00	-0.04**	-0.00	-0.06***
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains academic HS (age 31)	0.01	-0.05***	-0.00	-0.04	0.01	-0.02
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)
Attains short college (age 31)	0.00	-0.06***	0.00	-0.06	-0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.04)	(0.00)	(0.02)
Attains college (age 31)	0.00	-0.06***	0.00	-0.04	-0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.04)	(0.00)	(0.02)
Gross income (27-31 yrs old)	0.00	-151.06***	0.00	-289.69***	0.00	-179.16**
	(0.00)	(34.12)	(0.00)	(91.28)	(0.00)	(72.44)
Days/year unemp. (27-31 yrs old)	0.00	0.99	0.00	-3.24	0.00	4.88**
	(0.00)	(0.90)	(0.00)	(2.21)	(0.00)	(2.35)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Sample restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table C.3: Missingness in outcomes and covariates:
Between-school specification - 1977 cohort

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	0.01	-0.10***	-0.00	-0.06	0.01	-0.08**
	(0.02)	(0.02)	(0.01)	(0.04)	(0.03)	(0.03)
Advanced Math in grade 9	-0.03	-0.07***	-0.03	-0.06**	-0.04*	-0.06***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Advanced Eng in grade 9	-0.01	-0.05***	0.02	0.01	-0.04	-0.02
	(0.05)	(0.01)	(0.07)	(0.02)	(0.03)	(0.02)
Swedish grade (grade 9, 0-100)	0.01	-2.08***	-0.00	-1.08	0.01	-2.12***
	(0.02)	(0.38)	(0.01)	(0.92)	(0.03)	(0.73)
Math grade (grade 9, 0-100)	-0.00	-2.53***	-0.00	-3.21***	-0.00	-1.54**
	(0.01)	(0.44)	(0.01)	(1.01)	(0.01)	(0.71)
English grade (grade 9, 0-100)	-0.00	-1.30***	-0.00	1.11	-0.00	-0.88
	(0.01)	(0.39)	(0.01)	(0.93)	(0.01)	(0.78)
GPA (grade 9, 0-100)	0.00	-2.98***	0.00	-2.87***	0.00	-3.22***
		(0.38)		(0.89)		(0.63)
Long-term index (std munic)	0.00	-0.15***	-0.00	-0.18***	0.01	-0.15***
	(0.00)	(0.02)	(0.01)	(0.03)	(0.01)	(0.03)
Attains high school (age 31)	0.00	-0.05***	0.00	-0.04***	0.01	-0.06***
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Attains academic HS (age 31)	0.02	-0.06***	0.01	-0.02	0.03	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Attains short college (age 31)	0.00	-0.08***	0.00	-0.08***	0.01	-0.02
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.02)
Attains college (age 31)	0.00	-0.07***	0.00	-0.07***	0.01	-0.00
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Gross income (27-31 yrs old)	0.00	-217.76***	-0.00	-400.80***	0.01	-229.61***
	(0.00)	(29.87)	(0.01)	(61.78)	(0.01)	(49.40)
Days/year unemp. (27-31 yrs old)	0.00	2.86***	-0.00	2.32	0.01	6.34***
	(0.00)	(0.78)	(0.01)	(1.85)	(0.01)	(1.80)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether studying in a sampled school predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table C.4: Missingness in outcomes and covariates:
Within-municipality specification - 1982 cohort

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.01**	-0.10***	-0.00	-0.18***	-0.01*	0.01
	(0.00)	(0.02)	(0.00)	(0.05)	(0.01)	(0.03)
Swedish grade (grade 9, 0-100)	-0.01**	-2.26***	-0.00	-3.73***	-0.01*	-0.05
	(0.00)	(0.38)	(0.00)	(0.99)	(0.01)	(0.68)
Math grade (grade 9, 0-100)	-0.00	-2.86***	0.00	-5.07***	-0.00*	-0.78
	(0.00)	(0.42)	(0.00)	(0.92)	(0.00)	(0.61)
English grade (grade 9, 0-100)	-0.00	-1.99***	0.00	-1.70*	-0.00*	0.60
	(0.00)	(0.40)	(0.00)	(1.01)	(0.00)	(0.66)
GPA (grade 9, 0-100)	0.00	-3.83***	0.00	-5.93***	0.00	-2.70***
		(0.46)		(1.11)		(0.85)
Long-term index (std munic)	-0.00	-0.11***	-0.00	-0.21***	0.00	-0.12***
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.03)
Attains high school (age 31)	-0.00	-0.03***	-0.00	-0.04***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Attains academic HS (age 31)	-0.02**	-0.02**	-0.01	-0.01	-0.02*	0.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.03)
Attains short college (age 31)	-0.00	-0.05***	-0.00	-0.08***	0.00	-0.00
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains college (age 31)	-0.00	-0.04***	-0.00	-0.09***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Gross income (27-31 yrs old)	-0.00	-125.36***	-0.00	-339.54***	0.00	-215.04***
	(0.00)	(26.50)	(0.00)	(68.62)	(0.00)	(54.97)
Days/year unemp. (27-31 yrs old)	-0.00	4.06***	-0.00	6.69***	0.00	5.56***
	(0.00)	(0.74)	(0.00)	(1.84)	(0.00)	(1.65)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table C.5: Missingness in outcomes and covariates:
Within-municipality specification - 1982 cohort - Swedish born

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.00*	-0.07***	0.00	-0.12*	-0.01***	-0.05
	(0.00)	(0.02)	(0.00)	(0.06)	(0.00)	(0.05)
Swedish grade (grade 9, 0-100)	-0.00*	-1.19**	0.00	-1.79	-0.01***	-0.31
	(0.00)	(0.50)	(0.00)	(1.42)	(0.00)	(1.10)
Math grade (grade 9, 0-100)	-0.00**	-1.74***	0.00	-4.17***	-0.01**	-1.56
	(0.00)	(0.52)	(0.00)	(1.44)	(0.00)	(1.00)
English grade (grade 9, 0-100)	-0.00**	-0.69	0.00	0.63	-0.01**	0.65
	(0.00)	(0.51)	(0.00)	(1.52)	(0.00)	(1.13)
GPA (grade 9, 0-100)	0.00	-1.97***	0.00	-3.50**	0.00	-3.47***
		(0.50)		(1.37)		(1.31)
Long-term index (std munic)	0.00	-0.03*	0.00	-0.11**	0.00	-0.17***
	(0.00)	(0.02)	(0.00)	(0.05)	(0.00)	(0.05)
Attains high school (age 31)	0.00	-0.02**	0.00	-0.05**	0.00	-0.05*
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.03)
Attains academic HS (age 31)	-0.01	-0.02	-0.01	-0.03	-0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.04)
Attains short college (age 31)	0.00	-0.02*	0.00	-0.06	0.00	-0.07***
	(0.00)	(0.01)	(0.00)	(0.04)	(0.00)	(0.03)
Attains college (age 31)	0.00	-0.02	0.00	-0.09***	0.00	-0.07***
	(0.00)	(0.01)	(0.00)	(0.04)	(0.00)	(0.02)
Gross income (27-31 yrs old)	0.00	-16.05	0.00	-150.04	0.00	-235.54**
	(0.00)	(35.57)	(0.00)	(110.10)	(0.00)	(105.88)
Days/year unemp. (27-31 yrs old)	0.00	1.07	0.00	1.58	0.00	4.45*
	(0.00)	(0.82)	(0.00)	(2.19)	(0.00)	(2.40)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Sample restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table C.6: Missingness in outcomes and covariates:
Within-municipality specification - 1982 cohort - Same municipality

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.01	-0.21***	-0.00	-0.16***	-0.01	-0.01
	(0.00)	(0.03)	(0.00)	(0.05)	(0.01)	(0.04)
Swedish grade (grade 9, 0-100)	-0.01	-4.72***	-0.00	-3.16***	-0.01	-0.62
	(0.00)	(0.60)	(0.00)	(1.00)	(0.01)	(0.78)
Math grade (grade 9, 0-100)	-0.00	-5.66***	0.00	-4.80***	-0.00	-1.13*
	(0.00)	(0.59)	(0.00)	(0.97)	(0.00)	(0.67)
English grade (grade 9, 0-100)	-0.00	-3.93***	0.00	-1.18	-0.00	0.12
	(0.00)	(0.62)	(0.00)	(1.06)	(0.00)	(0.75)
GPA (grade 9, 0-100)	0.00	-7.87***	0.00	-5.62***	0.00	-3.30***
		(0.79)		(1.12)		(0.96)
Long-term index (std munic)	0.00	-0.27***	-0.00	-0.21***	0.00	-0.14***
	(0.00)	(0.02)	(0.00)	(0.04)	(0.00)	(0.03)
Attains high school (age 31)	0.00	-0.07***	-0.00	-0.05***	0.00	-0.04**
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains academic HS (age 31)	-0.01	-0.04**	-0.01	0.00	-0.00	0.03
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.03)
Attains short college (age 31)	0.00	-0.13***	-0.00	-0.08***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains college (age 31)	0.00	-0.13***	-0.00	-0.09***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Gross income (27-31 yrs old)	0.00	-349.33***	-0.00	-374.39***	0.00	-216.53***
	(0.00)	(46.22)	(0.00)	(71.75)	(0.00)	(61.56)
Days/year unemp. (27-31 yrs old)	0.00	8.65***	-0.00	5.87***	0.00	5.95***
	(0.00)	(1.21)	(0.00)	(1.95)	(0.00)	(1.75)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Sample restricted to students living in the same municipality in grades 3 and 9. Standard errors clustered at the school level.

Table C.7: Missingness in outcomes and covariates:
Between-school specification - 1982 cohort

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.02	-0.10***	-0.00	-0.18***	-0.04	0.01
	(0.01)	(0.02)	(0.00)	(0.05)	(0.02)	(0.03)
Swedish grade (grade 9, 0-100)	-0.02	-2.26***	-0.00	-3.73***	-0.04	-0.05
	(0.01)	(0.38)	(0.00)	(0.99)	(0.02)	(0.68)
Math grade (grade 9, 0-100)	-0.00	-2.86***	-0.00	-5.07***	-0.01	-0.78
	(0.00)	(0.42)	(0.00)	(0.92)	(0.00)	(0.61)
English grade (grade 9, 0-100)	-0.00	-1.99***	-0.00	-1.70*	-0.01	0.60
	(0.00)	(0.40)	(0.00)	(1.01)	(0.00)	(0.66)
GPA (grade 9, 0-100)	0.00	-3.83***	0.00	-5.93***	0.00	-2.70***
		(0.46)		(1.11)		(0.85)
Long-term index (std munic)	-0.00	-0.11***	-0.00	-0.21***	0.00	-0.12***
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.03)
Attains high school (age 31)	-0.00	-0.03***	-0.00	-0.04***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Attains academic HS (age 31)	-0.01	-0.02**	-0.01	-0.01	-0.02	0.03
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)
Attains short college (age 31)	-0.00	-0.05***	-0.00	-0.08***	0.00	-0.00
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains college (age 31)	-0.00	-0.04***	-0.00	-0.09***	0.00	-0.02
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Gross income (27-31 yrs old)	-0.00	-125.36***	-0.00	-339.54***	0.00	-215.04***
	(0.00)	(26.50)	(0.00)	(68.62)	(0.00)	(54.97)
Days/year unemp. (27-31 yrs old)	-0.00	4.06***	-0.00	6.69***	0.00	5.56***
	(0.00)	(0.74)	(0.00)	(1.84)	(0.00)	(1.65)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether studying in a sampled school predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table C.8: Missingness in outcomes and covariates:
Within-municipality specification - All

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.00	-0.10***	-0.00	-0.12***	-0.00	-0.04
	(0.00)	(0.01)	(0.00)	(0.03)	(0.01)	(0.02)
Swedish grade (grade 9, 0-100)	-0.00	-2.18***	-0.00	-2.50***	-0.00	-1.13**
	(0.00)	(0.28)	(0.00)	(0.68)	(0.01)	(0.50)
Math grade (grade 9, 0-100)	-0.00	-2.71***	-0.00	-4.22***	-0.00*	-1.16**
	(0.00)	(0.31)	(0.00)	(0.67)	(0.00)	(0.47)
English grade (grade 9, 0-100)	-0.00	-1.68***	-0.00	-0.39	-0.00	-0.15
	(0.00)	(0.29)	(0.00)	(0.69)	(0.00)	(0.50)
GPA (grade 9, 0-100)	0.00	-3.44***	0.00	-4.50***	0.00	-2.96***
		(0.33)		(0.71)		(0.58)
Long-term index (std munic)	0.00	-0.13***	0.00	-0.20***	0.00	-0.14***
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains high school (age 31)	0.00	-0.04***	0.00	-0.04***	0.00	-0.04***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Attains academic HS (age 31)	-0.01	-0.04***	-0.01	-0.01	-0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Attains short college (age 31)	0.00	-0.06***	0.00	-0.08***	0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Attains college (age 31)	0.00	-0.06***	0.00	-0.08***	0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Gross income (27-31 yrs old)	0.00	-166.97***	0.00	-368.09***	0.00	-221.80***
	(0.00)	(19.42)	(0.00)	(47.09)	(0.00)	(38.62)
Days/year unemp. (27-31 yrs old)	0.00	3.53***	0.00	4.67***	0.00	5.95***
	(0.00)	(0.55)	(0.00)	(1.26)	(0.00)	(1.26)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table C.9: Missingness in outcomes and covariates:
Within-municipality specification - All - Swedish born

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.00	-0.07***	0.00	-0.07	-0.00*	-0.08**
	(0.00)	(0.02)	(0.00)	(0.04)	(0.00)	(0.03)
Swedish grade (grade 9, 0-100)	-0.00*	-1.26***	0.00	-0.80	-0.01**	-1.58**
	(0.00)	(0.35)	(0.00)	(0.95)	(0.00)	(0.70)
Math grade (grade 9, 0-100)	-0.00*	-1.72***	-0.00	-2.82***	-0.00**	-1.34*
	(0.00)	(0.37)	(0.00)	(1.03)	(0.00)	(0.73)
English grade (grade 9, 0-100)	-0.00	-0.68**	0.00	0.37	-0.00*	-0.42
	(0.00)	(0.34)	(0.00)	(0.94)	(0.00)	(0.75)
GPA (grade 9, 0-100)	0.00	-1.86***	0.00	-2.80***	0.00	-3.26***
		(0.33)		(0.96)		(0.73)
Long-term index (std munic)	0.00*	-0.06***	0.00	-0.09***	0.00	-0.14***
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.03)
Attains high school (age 31)	0.00	-0.02***	0.00	-0.04***	0.00	-0.06***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.02)
Attains academic HS (age 31)	-0.00	-0.04***	-0.01	-0.03	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Attains short college (age 31)	0.00	-0.04***	0.00	-0.06**	0.00	-0.04**
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.02)
Attains college (age 31)	0.00	-0.04***	0.00	-0.07***	0.00	-0.03**
	(0.00)	(0.01)	(0.00)	(0.03)	(0.00)	(0.02)
Gross income (27-31 yrs old)	0.00	-81.51***	0.00	-226.86***	0.00	-202.09***
	(0.00)	(24.67)	(0.00)	(72.80)	(0.00)	(63.35)
Days/year unemp. (27-31 yrs old)	0.00	1.03*	0.00	-1.03	0.00	4.65***
	(0.00)	(0.60)	(0.00)	(1.51)	(0.00)	(1.65)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether treatment predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Sample restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

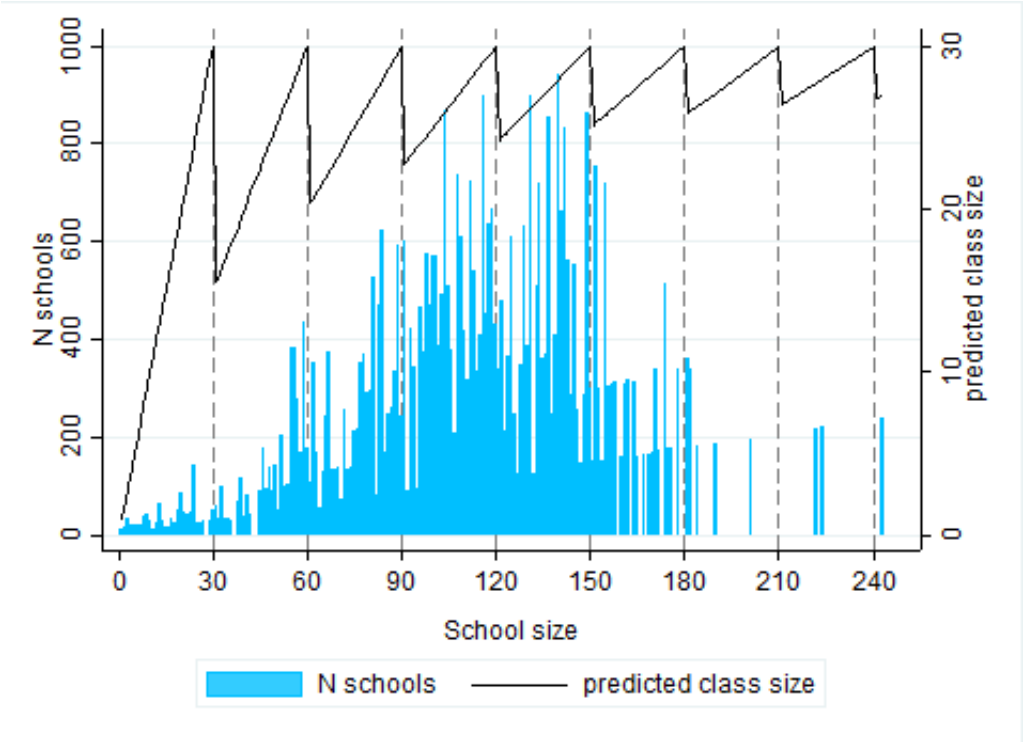
Table C.10: Missingness in outcomes and covariates:
Between-school specification - All

Sample	All		High Education		Low Education	
Test for missing	Y	X	Y	X	Y	X
Short-term index (std munic)	-0.01	-0.10***	-0.00	-0.12***	-0.02	-0.04
	(0.01)	(0.01)	(0.00)	(0.03)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	-0.01	-2.18***	-0.00	-2.50***	-0.02	-1.13**
	(0.01)	(0.28)	(0.00)	(0.68)	(0.02)	(0.50)
Math grade (grade 9, 0-100)	-0.00	-2.71***	-0.00	-4.22***	-0.01	-1.16**
	(0.00)	(0.31)	(0.00)	(0.67)	(0.00)	(0.47)
English grade (grade 9, 0-100)	-0.00	-1.68***	-0.00	-0.39	-0.00	-0.15
	(0.00)	(0.29)	(0.00)	(0.69)	(0.00)	(0.50)
GPA (grade 9, 0-100)	0.00	-3.44***	0.00	-4.50***	0.00	-2.96***
		(0.33)		(0.71)		(0.58)
Long-term index (std munic)	0.00	-0.13***	-0.00	-0.20***	0.01	-0.14***
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.02)
Attains high school (age 31)	-0.00	-0.04***	-0.00	-0.04***	0.00	-0.04***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Attains academic HS (age 31)	-0.00	-0.04***	-0.00	-0.01	-0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Attains short college (age 31)	-0.00	-0.06***	-0.00	-0.08***	0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Attains college (age 31)	-0.00	-0.06***	-0.00	-0.08***	0.00	-0.01
	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.01)
Gross income (27-31 yrs old)	0.00	-166.97***	-0.00	-368.09***	0.01*	-221.80***
	(0.00)	(19.42)	(0.00)	(47.09)	(0.00)	(38.62)
Days/year unemp. (27-31 yrs old)	0.00	3.53***	-0.00	4.67***	0.01*	5.95***
	(0.00)	(0.55)	(0.00)	(1.26)	(0.00)	(1.26)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents two sets of tests, for three samples: all students, students from highly educated families, and students from families with low education. Column Y tests whether studying in a sampled school predicts missingness of the outcome variable. Column X tests whether observing all covariates predict the level of the outcome variable. Each regression controls for school size and municipality fixed effects. The main outcomes are standardized indices for short-term educational attainment and long-term educational attainment and labor market performance. Individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Figure D.1: Distribution of School Size and Predicted Class Size.

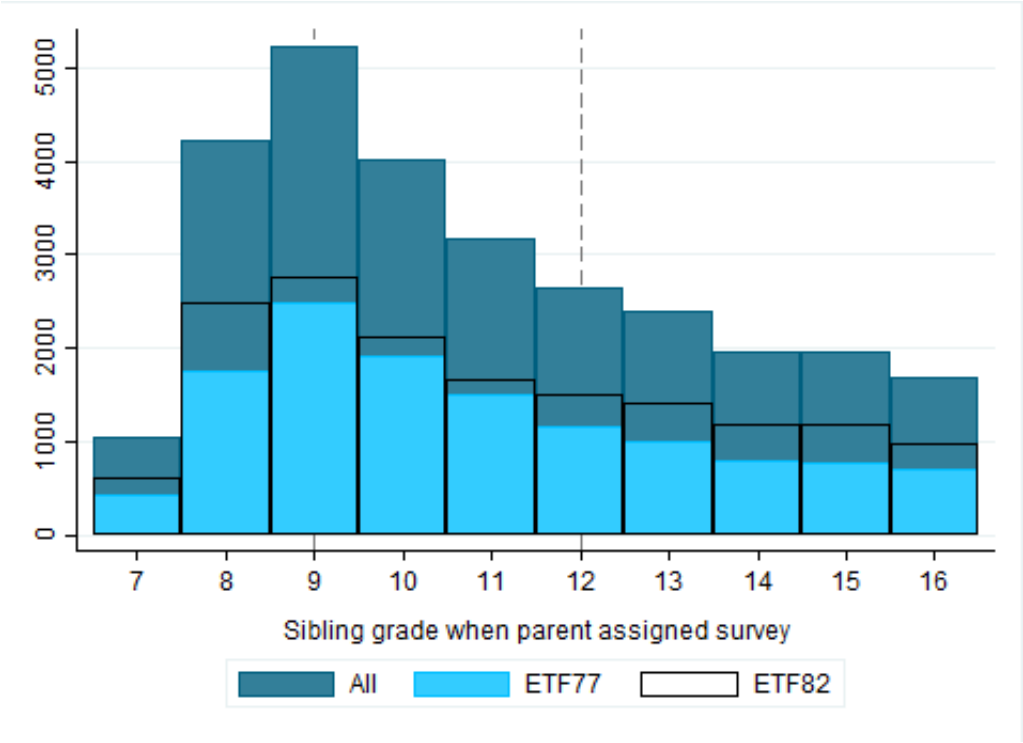


The Figure displays the number of schools (left y-axis) and predicted class size according to Maimonides' rule (right y-axis) by school size (x-axis) of the focal child.

D Descriptive Statistics

D.1 Figures

Figure D.2: Distribution of Siblings, by Grade and Cohort.



The Figure displays the number of older siblings (y-axis) by stipulated grade of the older sibling when their parent was surveyed (x-axis) and the focal child attended 6th grade. The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. The number of older siblings is displayed separately by cohort.

D.2 Outcome Variables

Table D.1: Outcome means: 1977 cohort

Sample	All	High Education	Low Education
Short-term index (std munic)	0.05	0.25	-0.22
Advanced Math in grade 9	0.62	0.74	0.45
Advanced Eng in grade 9	0.76	0.86	0.63
Swedish grade (grade 9, 0-100)	64.92	69.12	59.03
Math grade (grade 9, 0-100)	63.47	66.58	59.25
English grade (grade 9, 0-100)	65.50	68.74	61.08
GPA (grade 9, 0-100)	64.76	68.42	59.78
Long-term index (std munic)	0.03	0.17	-0.15
Attains high school (age 31)	0.93	0.95	0.88
Attains academic HS (age 31)	0.58	0.71	0.39
Attains short college (age 31)	0.49	0.62	0.30
Attains college (age 31)	0.36	0.48	0.20
Gross income (27-31 yrs old)	2,124.75	2,226.38	1,987.07
Days/year unemp. (27-31 yrs old)	16.49	14.16	19.65
Observations	17066	9947	7119
Schools	221	220	219

The table presents mean outcome values for the 1977 cohort. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13.

Table D.2: Variation in outcomes: 1977 cohort

Level of variation	Overall	Municipality		School	
		Between	Within	Between	Within
Short-term index (std munic)	0.80	0.06	0.80	0.22	0.78
Advanced Math in grade 9	0.49	0.07	0.48	0.16	0.47
Advanced Eng in grade 9	0.43	0.05	0.42	0.15	0.41
Swedish grade (grade 9, 0-100)	17.67	2.52	17.57	4.76	17.21
Math grade (grade 9, 0-100)	18.51	2.13	18.45	5.10	18.10
English grade (grade 9, 0-100)	17.58	1.85	17.49	4.84	17.14
GPA (grade 9, 0-100)	15.30	1.65	15.25	14.51	13.76
Long-term index (std munic)	0.67	0.03	0.67	0.29	0.66
Attains high school (age 31)	0.26	0.03	0.26	0.09	0.26
Attains academic HS (age 31)	0.49	0.09	0.48	0.18	0.47
Attains short college (age 31)	0.50	0.06	0.50	0.17	0.48
Attains college (age 31)	0.48	0.05	0.48	0.14	0.47
Gross income (27-31 yrs old)	1,313.28	129.77	1,308.41	346.70	1,292.41
Days/year unemp. (27-31 yrs old)	35.21	4.63	34.99	18.28	34.53
Observations	17066	17066	17066	17066	17066
Schools	221	221	221	221	221

The table presents the amount of variation in outcomes for the 1977 cohort. Column 1 considers all sample. Columns 2 and 3 show the standard deviation of each outcome between and within municipality. Columns 4 and 5 the standard deviation of each outcome between and within school. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13.

Table D.3: Outcome means: 1982 cohort

Sample	All	High Education	Low Education
Short-term index (std munic)	0.07	0.28	-0.27
Swedish grade (grade 9, 0-100)	66.43	70.73	59.38
Math grade (grade 9, 0-100)	62.08	66.45	55.18
English grade (grade 9, 0-100)	66.80	71.27	59.74
GPA (grade 9, 0-100)	64.04	68.81	56.56
Long-term index (std munic)	0.05	0.17	-0.14
Attains high school (age 31)	0.91	0.95	0.85
Attains academic HS (age 31)	0.73	0.83	0.54
Attains short college (age 31)	0.56	0.69	0.35
Attains college (age 31)	0.40	0.51	0.22
Gross income (27-31 yrs old)	2,206.07	2,297.37	2,064.11
Days/year unemp. (27-31 yrs old)	13.98	11.38	18.02
Observations	21813	13530	8283
Schools	343	332	332

The table presents mean outcome values for the 1982 cohort. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13.

Table D.4: Variation in outcomes: 1982 cohort

Level of variation	Overall	Municipality		School	
		Between	Within	Between	Within
Short-term index (std munic)	0.85	0.15	0.85	0.59	0.80
Swedish grade (grade 9, 0-100)	19.27	2.53	19.13	11.41	18.10
Math grade (grade 9, 0-100)	18.89	2.85	18.77	10.79	17.90
English grade (grade 9, 0-100)	19.51	2.25	19.40	11.79	18.27
GPA (grade 9, 0-100)	18.87	2.80	18.74	15.90	17.05
Long-term index (std munic)	0.68	0.07	0.68	0.29	0.66
Attains high school (age 31)	0.28	0.04	0.28	0.17	0.28
Attains academic HS (age 31)	0.44	0.11	0.43	0.19	0.42
Attains short college (age 31)	0.50	0.09	0.49	0.21	0.48
Attains college (age 31)	0.49	0.09	0.49	0.18	0.47
Gross income (27-31 yrs old)	1,379.00	108.15	1,375.31	444.63	1,353.16
Days/year unemp. (27-31 yrs old)	33.53	3.20	33.36	11.29	32.93
Observations	21813	21813	21813	21813	21813
Schools	343	343	343	343	343

The table presents the amount of variation in outcomes for the 1982 cohort. Column 1 considers all sample. Columns 2 and 3 show the standard deviation of each outcome between and within municipality. Columns 4 and 5 the standard deviation of each outcome between and within school. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13.

Table D.5: Outcome means: All

Sample	All	High Education	Low Education
Short-term index (std munic)	0.07	0.27	-0.24
Swedish grade (grade 9, 0-100)	65.77	70.05	59.22
Math grade (grade 9, 0-100)	62.69	66.51	57.05
English grade (grade 9, 0-100)	66.23	70.20	60.35
GPA (grade 9, 0-100)	64.36	68.64	58.04
Long-term index (std munic)	0.04	0.17	-0.14
Attains high school (age 31)	0.92	0.95	0.87
Attains academic HS (age 31)	0.66	0.78	0.46
Attains short college (age 31)	0.53	0.66	0.33
Attains college (age 31)	0.38	0.50	0.21
Gross income (27-31 yrs old)	2,170.31	2,267.13	2,028.67
Days/year unemp. (27-31 yrs old)	15.09	12.57	18.77
Observations	38879	23477	15402
Schools	471	459	460

The table presents mean outcome values for the full sample. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13.

Table D.6: Variation in outcomes: All

Level of variation	Overall	Municipality		School	
		Between	Within	Between	Within
Short-term index (std munic)	0.83	0.12	0.83	0.51	0.79
Swedish grade (grade 9, 0-100)	18.60	2.64	18.47	9.71	17.78
Math grade (grade 9, 0-100)	18.74	2.51	18.65	9.40	18.05
English grade (grade 9, 0-100)	18.70	2.07	18.60	10.04	17.85
GPA (grade 9, 0-100)	17.39	2.17	17.31	14.27	15.77
Long-term index (std munic)	0.68	0.06	0.68	0.28	0.66
Attains high school (age 31)	0.28	0.03	0.27	0.15	0.27
Attains academic HS (age 31)	0.47	0.12	0.46	0.19	0.45
Attains short college (age 31)	0.50	0.08	0.50	0.19	0.48
Attains college (age 31)	0.49	0.07	0.48	0.16	0.47
Gross income (27-31 yrs old)	1,351.08	118.13	1,346.96	402.25	1,329.39
Days/year unemp. (27-31 yrs old)	34.30	3.83	34.10	12.60	33.71
Observations	38879	38879	38879	38879	38879
Schools	471	471	471	471	471

The table presents the amount of variation in outcomes for the full sample. Column 1 considers the full sample. Columns 2 and 3 show the standard deviation of each outcome between and within municipality. Columns 4 and 5 the standard deviation of each outcome between and within school. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13.

E Empirical Results

Table E.1: Effect of being surveyed: Within-municipality specification
1977 cohort

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Short-term index (std munic)	-0.02	-0.02	0.01	0.00	-0.04*	-0.05*
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Advanced Math in grade 9	-0.01	-0.01	-0.01	-0.01	-0.02	-0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Advanced Eng in grade 9	-0.01	-0.01	-0.01	-0.01	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	-0.18	-0.21	0.35	0.16	-0.66	-0.84*
	(0.44)	(0.36)	(0.54)	(0.49)	(0.51)	(0.50)
Math grade (grade 9, 0-100)	-0.35	-0.28	0.31	0.19	-0.86	-0.95
	(0.44)	(0.39)	(0.57)	(0.55)	(0.57)	(0.58)
English grade (grade 9, 0-100)	-0.51	-0.47	-0.13	-0.19	-0.69	-0.70
	(0.43)	(0.40)	(0.55)	(0.52)	(0.54)	(0.56)
GPA (grade 9, 0-100)	-0.34	-0.33	0.06	-0.14	-0.65	-0.79*
	(0.39)	(0.31)	(0.48)	(0.44)	(0.43)	(0.40)
Long-term index (std munic)	0.00	0.01	0.01	0.01	-0.00	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains high school (age 31)	-0.00	-0.00	-0.00	-0.01	0.00	0.00
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Attains academic HS (age 31)	-0.02	-0.02*	-0.01	-0.01	-0.04**	-0.04**
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains short college (age 31)	-0.02	-0.01*	-0.01	-0.01	-0.02	-0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains college (age 31)	-0.01	-0.01	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Gross income (27-31 yrs old)	-3.75	8.20	20.37	29.40	-25.21	-25.25
	(26.59)	(24.70)	(37.49)	(38.48)	(39.05)	(36.77)
Days/year unemp. (27-31 yrs old)	-1.04	-1.06	-0.68	-0.57	-1.49	-1.41
	(0.69)	(0.69)	(0.84)	(0.85)	(1.18)	(1.19)
Observations	17066	17066	9947	9947	7119	7119
Schools	221	221	220	220	219	219

* p < 0.10, ** p < 0.05, *** p < 0.01

The table presents coefficient estimates and standard errors from the regression of each outcome (in rows) on an indicator for being surveyed; i.e. δ_1 in (1). All specifications control for school size and municipality fixed effects. Additional controls include family background and municipality of birth. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table E.2: Effect of being surveyed: Within-municipality specification
1977 cohort - Swedish born

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Short-term index (std munic)	-0.02	-0.02	0.00	-0.00	-0.04	-0.05*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Advanced Math in grade 9	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Advanced Eng in grade 9	-0.01	-0.01	-0.01	-0.01	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	-0.35	-0.50	0.16	-0.09	-0.78	-1.05*
	(0.49)	(0.41)	(0.59)	(0.54)	(0.58)	(0.55)
Math grade (grade 9, 0-100)	-0.35	-0.36	0.19	0.14	-0.75	-0.88
	(0.49)	(0.44)	(0.64)	(0.60)	(0.64)	(0.64)
English grade (grade 9, 0-100)	-0.43	-0.49	-0.02	-0.12	-0.65	-0.76
	(0.49)	(0.45)	(0.59)	(0.57)	(0.63)	(0.63)
GPA (grade 9, 0-100)	-0.38	-0.47	-0.16	-0.32	-0.50	-0.74*
	(0.44)	(0.36)	(0.52)	(0.47)	(0.48)	(0.44)
Long-term index (std munic)	-0.01	-0.01	-0.00	-0.00	-0.02	-0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains high school (age 31)	-0.01	-0.01	-0.01	-0.01	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains academic HS (age 31)	-0.02	-0.02	-0.02	-0.02	-0.03	-0.03
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains short college (age 31)	-0.02	-0.02**	-0.02	-0.02	-0.02	-0.03*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains college (age 31)	-0.01	-0.02	-0.02	-0.02	-0.01	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Gross income (27-31 yrs old)	-27.48	-6.32	-4.77	15.68	-47.07	-32.64
	(29.09)	(27.51)	(42.36)	(42.75)	(41.12)	(37.45)
Days/year unemp. (27-31 yrs old)	-0.63	-0.60	-0.67	-0.52	-0.64	-0.58
	(0.74)	(0.76)	(0.86)	(0.88)	(1.27)	(1.31)
Observations	13817	13817	8130	8130	5687	5687
Schools	220	220	219	219	217	217

* p < 0.10, ** p < 0.05, *** p < 0.01

The table presents coefficient estimates and standard errors from the regression of each outcome (in rows) on an indicator for being surveyed; i.e. δ_1 in (1). All specifications control for school size and municipality fixed effects. Additional controls include family background and municipality of birth. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table E.3: Effect of assignment to sampled school: 1977 cohort

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Short-term index (std munic)	-0.03	-0.01	-0.02	-0.01	-0.03	-0.02
	(0.04)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)
Advanced Math in grade 9	-0.02	-0.01	-0.02	-0.01	-0.02	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Advanced Eng in grade 9	-0.01	-0.01	0.01	0.01	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	-0.71	-0.36	-0.35	-0.08	-1.10*	-0.77
	(0.85)	(0.55)	(0.81)	(0.60)	(0.64)	(0.68)
Math grade (grade 9, 0-100)	-1.20	-0.68	-0.65	-0.42	-1.54*	-1.08
	(0.86)	(0.58)	(0.80)	(0.69)	(0.81)	(0.79)
English grade (grade 9, 0-100)	-0.16	0.02	-0.47	-0.26	0.47	0.56
	(0.83)	(0.56)	(0.81)	(0.65)	(0.75)	(0.76)
GPA (grade 9, 0-100)	-0.31	0.07	-0.18	0.06	-0.18	0.14
	(0.80)	(0.59)	(0.82)	(0.73)	(0.55)	(0.58)
Long-term index (std munic)	-0.05	-0.03*	-0.03	-0.02	-0.06**	-0.03
	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Attains high school (age 31)	-0.02**	-0.01**	-0.01	-0.01	-0.03**	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains academic HS (age 31)	0.01	0.01	0.01	0.01	0.01	0.01
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Attains short college (age 31)	-0.02	-0.01	-0.00	0.00	-0.03*	-0.02
	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains college (age 31)	-0.01	0.00	-0.01	-0.00	-0.00	0.00
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Gross income (27-31 yrs old)	-73.03*	-52.47*	-52.03	-29.50	-95.83**	-79.71*
	(37.74)	(27.09)	(47.11)	(34.58)	(43.55)	(42.64)
Days/year unemp. (27-31 yrs old)	1.49	0.78	1.03	0.65	1.97	0.45
	(1.08)	(0.86)	(1.07)	(1.06)	(1.55)	(1.42)
Observations	17066	17066	9947	9947	7119	7119
Schools	221	221	220	220	219	219

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents coefficient estimates and standard errors from the regression of each outcome (in rows) on an indicator for studying in a sampled school; i.e. α_1 in (2). All specifications control for school size and municipality fixed effects. Additional controls include family background and municipality of birth. The main outcomes are standardized indices for short-term educational achievement, and long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table E.4: Effect of being surveyed: Within-municipality specification

All

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Short-term index (std munic)	0.02	0.01	0.02	0.02	-0.01	-0.02
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	0.11	-0.08	0.24	0.17	-0.37	-0.53
	(0.29)	(0.25)	(0.36)	(0.33)	(0.35)	(0.33)
Math grade (grade 9, 0-100)	0.38	0.21	0.59	0.52	-0.18	-0.36
	(0.29)	(0.26)	(0.36)	(0.35)	(0.36)	(0.35)
English grade (grade 9, 0-100)	0.10	-0.03	0.21	0.22	-0.29	-0.39
	(0.29)	(0.27)	(0.36)	(0.35)	(0.35)	(0.35)
GPA (grade 9, 0-100)	0.53*	0.29	0.50	0.42	0.21	-0.07
	(0.29)	(0.25)	(0.33)	(0.31)	(0.34)	(0.32)
Long-term index (std munic)	0.03**	0.02*	0.02	0.02	0.03*	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains high school (age 31)	0.00	-0.00	0.00	-0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Attains academic HS (age 31)	-0.01	-0.01	-0.01	-0.01	-0.02	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains short college (age 31)	0.00	-0.00	-0.01	-0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains college (age 31)	0.00	0.00	0.00	0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Gross income (27-31 yrs old)	37.06**	26.72	49.64*	50.41**	18.67	-5.85
	(17.90)	(16.36)	(25.48)	(24.28)	(24.30)	(23.50)
Days/year unemp. (27-31 yrs old)	-0.99**	-0.69	-0.49	-0.36	-1.56**	-1.17
	(0.46)	(0.44)	(0.51)	(0.51)	(0.79)	(0.79)
Observations	38879	38879	23477	23477	15402	15402
Schools	471	471	459	459	460	460

* p < 0.10, ** p < 0.05, *** p < 0.01

The table presents coefficient estimates and standard errors from the regression of each outcome (in rows) on an indicator for being surveyed; i.e. δ_1 in (1). All specifications control for school size and municipality fixed effects. Additional controls include family background and municipality of birth. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table E.5: Effect of being surveyed: Within-municipality specification
All - Swedish born

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Short-term index (std munic)	0.01	0.01	0.02	0.02	-0.01	-0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	-0.02	-0.12	0.13	0.16	-0.43	-0.60
	(0.31)	(0.26)	(0.37)	(0.34)	(0.39)	(0.36)
Math grade (grade 9, 0-100)	0.26	0.21	0.54	0.62	-0.28	-0.54
	(0.32)	(0.29)	(0.39)	(0.39)	(0.41)	(0.41)
English grade (grade 9, 0-100)	0.07	0.01	0.22	0.28	-0.25	-0.39
	(0.32)	(0.29)	(0.39)	(0.37)	(0.40)	(0.39)
GPA (grade 9, 0-100)	0.36	0.23	0.38	0.41	0.10	-0.18
	(0.31)	(0.26)	(0.33)	(0.31)	(0.38)	(0.36)
Long-term index (std munic)	0.02	0.01	0.02	0.02*	0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains high school (age 31)	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Attains academic HS (age 31)	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains short college (age 31)	-0.00	-0.00	-0.01	-0.01	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains college (age 31)	-0.00	-0.00	-0.01	-0.00	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Gross income (27-31 yrs old)	19.24	17.76	40.80	50.53*	-8.97	-27.52
	(19.84)	(18.43)	(28.57)	(27.32)	(25.59)	(24.53)
Days/year unemp. (27-31 yrs old)	-0.71	-0.61	-0.66	-0.65	-0.79	-0.56
	(0.45)	(0.45)	(0.49)	(0.50)	(0.84)	(0.84)
Observations	30915	30915	18852	18852	12063	12063
Schools	463	463	452	452	441	441

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents coefficient estimates and standard errors from the regression of each outcome (in rows) on an indicator for being surveyed; i.e. δ_1 in (1). All specifications control for school size and municipality fixed effects. Additional controls include family background and municipality of birth. The main outcomes are standardized indices for short-term educational attainment, and long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table E.6: Effect of assignment to sampled school
All

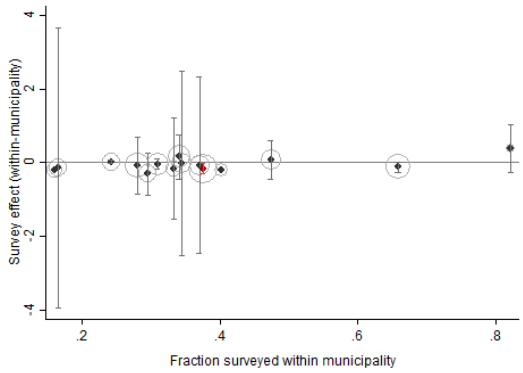
Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Short-term index (std munic)	0.02	0.02	0.04	0.04	-0.02	-0.02
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Swedish grade (grade 9, 0-100)	0.22	0.15	0.60	0.57	-0.64	-0.66
	(0.55)	(0.45)	(0.63)	(0.57)	(0.44)	(0.42)
Math grade (grade 9, 0-100)	0.59	0.52	0.85	0.87	-0.10	-0.20
	(0.54)	(0.41)	(0.58)	(0.54)	(0.46)	(0.44)
English grade (grade 9, 0-100)	1.18**	1.08**	1.21*	1.23**	0.77*	0.73
	(0.59)	(0.45)	(0.65)	(0.58)	(0.47)	(0.45)
GPA (grade 9, 0-100)	0.65	0.54	0.75	0.75	0.04	0.00
	(0.60)	(0.47)	(0.67)	(0.61)	(0.46)	(0.44)
Long-term index (std munic)	0.01	0.00	-0.00	0.00	0.00	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Attains high school (age 31)	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Attains academic HS (age 31)	0.00	-0.00	-0.01	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Attains short college (age 31)	0.01	0.00	0.00	0.00	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains college (age 31)	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Gross income (27-31 yrs old)	18.03	16.13	7.37	15.38	23.73	8.35
	(24.67)	(20.40)	(29.92)	(26.55)	(30.00)	(27.44)
Days/year unemp. (27-31 yrs old)	0.13	0.23	0.36	0.39	0.12	0.23
	(0.58)	(0.45)	(0.52)	(0.50)	(0.98)	(0.85)
Observations	38879	38879	23477	23477	15402	15402
Schools	471	471	459	459	460	460

* p < 0.10, ** p < 0.05, *** p < 0.01

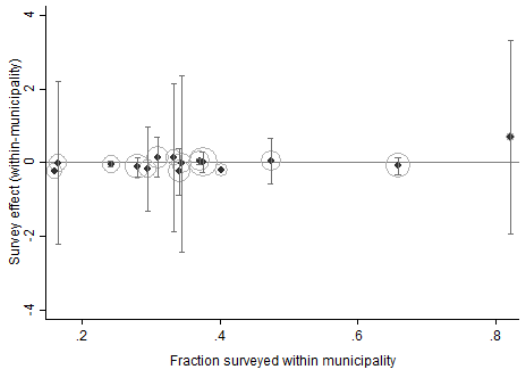
The table presents coefficient estimates and standard errors from the regression of each outcome (in rows) on an indicator for studying in a sampled school; i.e. α_1 in (2). All specifications control for school size and municipality fixed effects. Additional controls include family background and municipality of birth. The main outcomes are standardized indices for short-term educational achievement, and long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

F Heterogeneity by Municipality

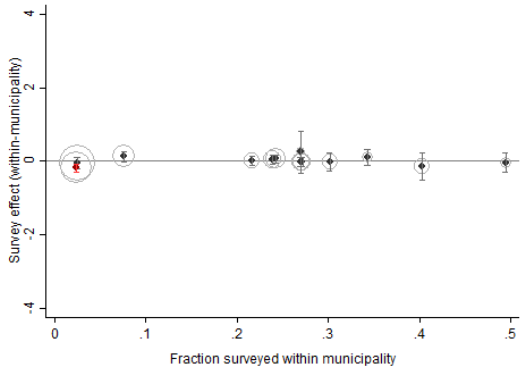
Figure F.1: Survey Effect, by Municipality - 1977 cohort



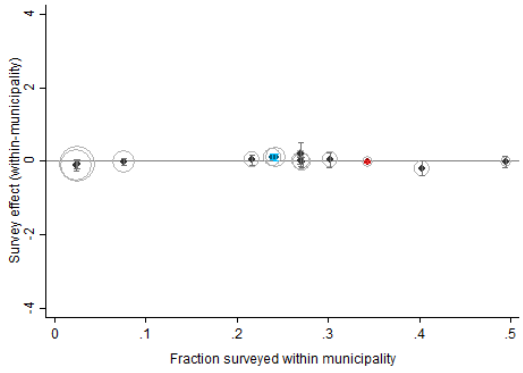
(a) Short-term Index
Within-municipality, All schools surveyed.



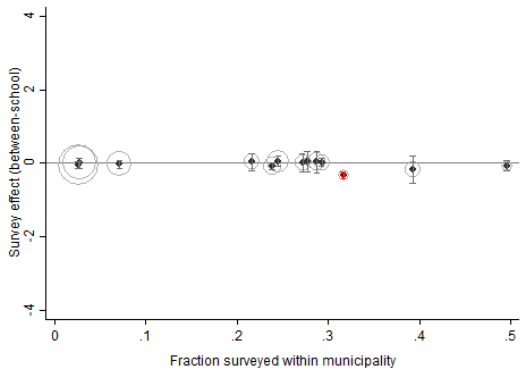
(b) Long-term Index
Within-municipality, All schools surveyed.



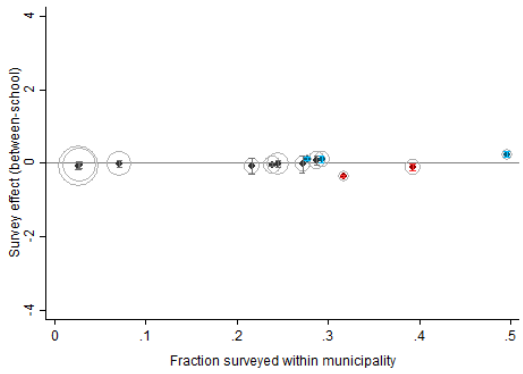
(c) Short-term Index
Within-municipality, Some schools surveyed.



(d) Long-term Index
Within-municipality, Some schools surveyed.



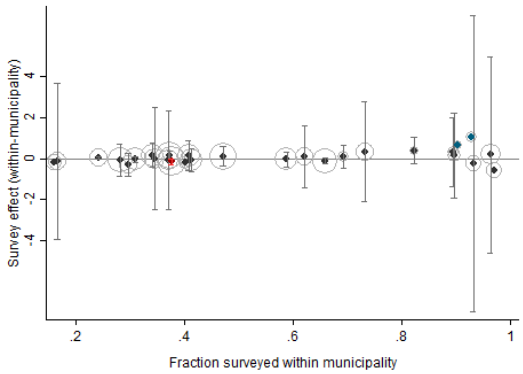
(e) Short-term Index
Between-school, Some schools surveyed.



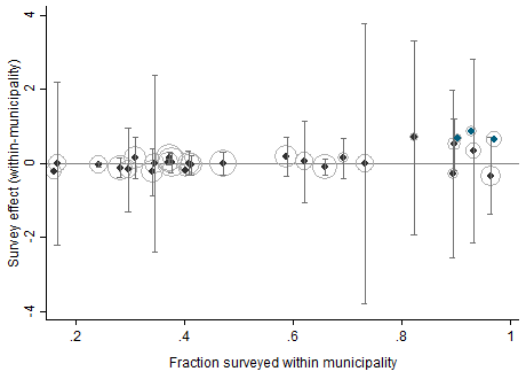
(f) Long-term Index
Between-school, Some schools surveyed.

The Figure displays the survey effect on the short- and long-term index (y-axis) by fraction treated within the municipality (x-axis). The outcome variable in the first column is the short-term index, and the outcome in the second column is the long-term index. The sample is split by whether all schools in the municipality have sampled classes (first row) or not (last two rows). The first two rows show the survey effect δ_1 in (1) by municipality, while the last row shows the between-school estimate of α_1 (the coefficient to the $ETFSchool$ indicator) in (2). All estimates are shown by municipality instead of adding municipality fixed effects. Controls include school size, family background, and municipality of birth. Solid diamonds mark point estimates and I-beams mark 95% confidence intervals. Significantly positive estimates are marked with blue, while significantly negative estimates are marked with red. Hollow circles represent municipality size in terms of the number of students. Sample: ETF77 cohort. Swedish born for within-municipality specification and full sample for between-school specification.

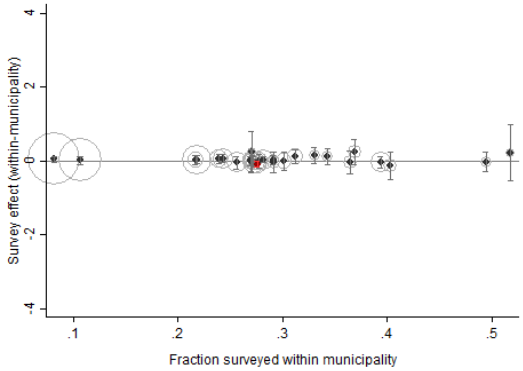
Figure F.2: Survey Effect, by Municipality - All



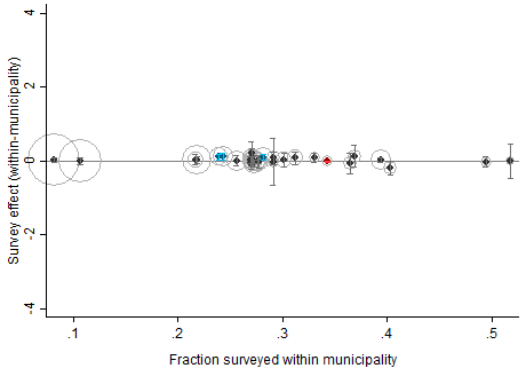
(a) Short-term Index
Within-municipality, All schools surveyed.



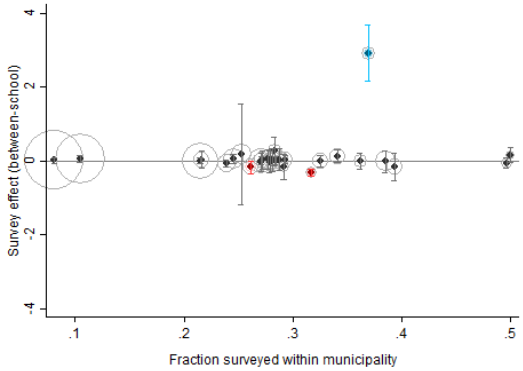
(b) Long-term Index
Within-municipality, All schools surveyed.



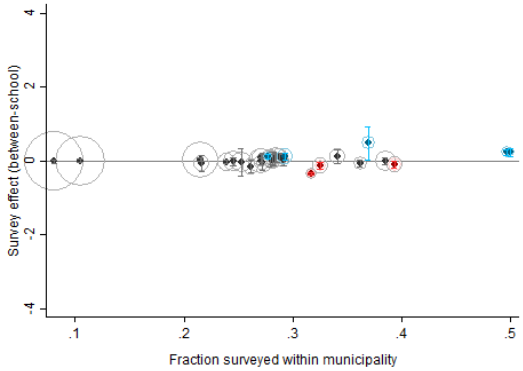
(c) Short-term Index
Within-municipality, Some schools surveyed.



(d) Long-term Index
Within-municipality, Some schools surveyed.



(e) Short-term Index
Between-school, Some schools surveyed.

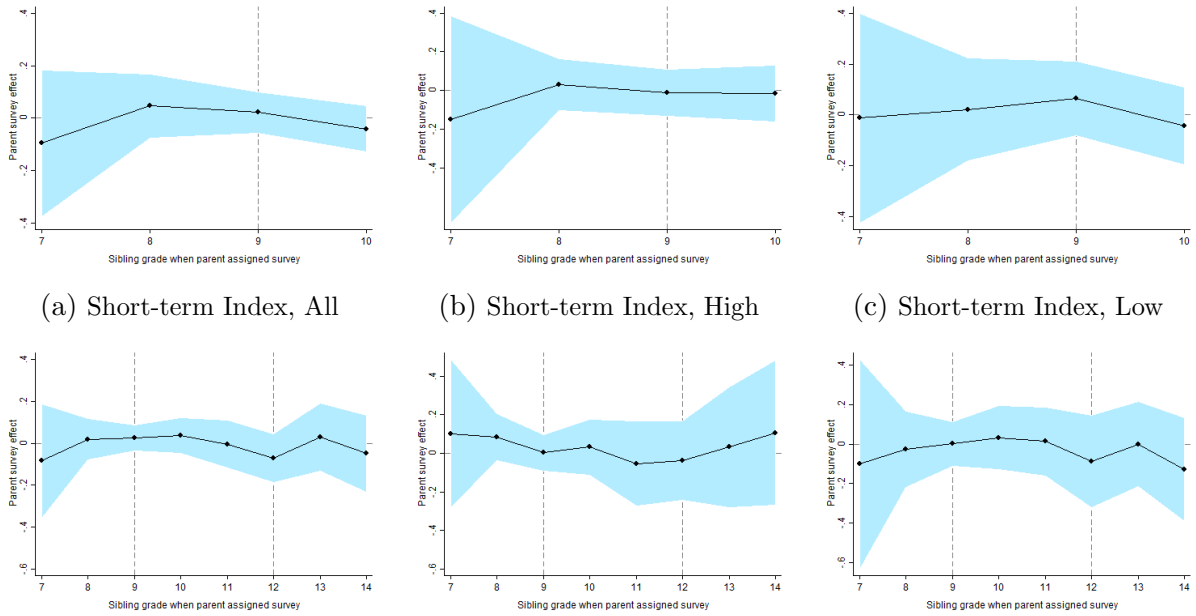


(f) Long-term Index
Between-school, Some schools surveyed.

The Figure displays the survey effect on the short- and long-term index (y-axis) by fraction treated within the municipality (x-axis). The outcome variable in the first column is the short-term index, and the outcome in the second column is the long-term index. The sample is split by whether all schools in the municipality have sampled classes (first row) or not (last two rows). The first two rows show the survey effect δ_1 in (1) by municipality, while the last row shows the between-school estimate of α_1 (the coefficient to the $ETFSchool$ indicator) in (2). All estimates are shown by municipality instead of adding municipality fixed effects. Controls include school size, family background, and municipality of birth. Solid diamonds mark point estimates and I-beams mark 95% confidence intervals. Significantly positive estimates are marked with blue, while significantly negative estimates are marked with red. Hollow circles represent municipality size in terms of the number of students. Sample: Swedish born for within-municipality specification and full sample for between-school specification.

G Sibling Spillovers

Figure G.1: Short- and Long-Term Effect of Parent Survey on Siblings, Within-Municipality - 1977 Cohort.



(a) Short-term Index, All (b) Short-term Index, High (c) Short-term Index, Low

(d) Long-term Index, All (e) Long-term Index, High (f) Long-term Index, Low

The Figure displays the parental survey effect on the short- and long-term index of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: ETF77 cohort, Swedish born.

Table G.1: Effect of parent being surveyed: Within-municipality specification
Siblings of 1977 cohort - Swedish born

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Long-term index (std munic)	0.00	0.01	-0.02	-0.01	0.02	0.00
	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.04)
Attains high school (age 31)	-0.01	-0.01	-0.00	-0.00	-0.02	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Attains academic HS (age 31)	-0.03	-0.03	-0.05*	-0.06**	0.01	0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Attains short college (age 31)	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Attains college (age 31)	-0.03	-0.03*	-0.03	-0.03	-0.02	-0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Gross income (27-31 yrs old)	29.74	46.48	22.36	59.30	33.25	1.16
	(50.97)	(47.11)	(79.82)	(76.53)	(68.83)	(66.81)
Days/year unemp. (27-31 yrs old)	-1.66	-1.31	0.52	0.55	-3.36	-2.26
	(1.38)	(1.46)	(1.78)	(1.80)	(2.27)	(2.38)
Observations	6537	6537	3930	3930	2607	2607
Schools	226	226	222	222	216	216

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents coefficient estimates and standard errors from the regression (1) of each outcome (in rows) on an indicator for focal child being surveyed interacted with sibling being in the critical grades after compulsory schooling, $g \in \{10, 11, 12\}$, controlling for municipality fixed effects, sibling grade, school size, and the indicator for focal child being surveyed interacted with siblings being in earlier and later grades, $g \in \{7, 8, 9\}$ and $g \in \{13, 14, 15, 16\}$, respectively. Additional controls include family background and municipality of birth. The main outcome is the standardized index for long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table G.2: Effect of sibling at sampled school: Siblings of 1977 cohort

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Long-term index (std munic)	0.00	0.01	0.05	0.05	-0.07	-0.06
	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)
Attains high school (age 31)	0.00	0.01	0.02*	0.02*	-0.02	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Attains academic HS (age 31)	-0.02	-0.03	-0.02	-0.04	-0.01	-0.00
	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.03)
Attains short college (age 31)	-0.01	-0.01	-0.00	-0.02	-0.01	0.00
	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)
Attains college (age 31)	-0.00	-0.00	-0.00	-0.02	0.01	0.03
	(0.02)	(0.02)	(0.04)	(0.03)	(0.02)	(0.02)
Gross income (27-31 yrs old)	-35.53	0.26	68.37	85.46	-157.10**	-144.35**
	(70.85)	(58.22)	(92.26)	(80.31)	(70.67)	(68.08)
Days/year unemp. (27-31 yrs old)	-1.93	-1.95	-4.30**	-4.63**	2.57	3.06
	(1.58)	(1.48)	(1.82)	(1.82)	(2.35)	(2.37)
Observations	6537	6537	3930	3930	2607	2607
Schools	226	226	222	222	216	216

* p < 0.10, ** p < 0.05, *** p < 0.01

The table presents coefficient estimates and standard errors from the regression (2) of each outcome (in rows) on an indicator for focal child being in a sampled school interacted with sibling being in the critical grades after compulsory schooling, $g \in \{10, 11, 12\}$, controlling for municipality fixed effects, sibling grade, school size, and the indicator for focal child being surveyed interacted with siblings being in earlier and later grades, $g \in \{7, 8, 9\}$ and $g \in \{13, 14, 15, 16\}$, respectively. Additional controls include family background and municipality of birth. The main outcome is the standardized index for long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Table G.3: Effect of parent being surveyed: Within-municipality specification
All Siblings of 1977 and 1982 cohort - Swedish born

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Long-term index (std munic)	0.05***	0.04**	0.03	0.03	0.05	0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Attains high school (age 31)	0.01	0.01	0.00	0.00	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Attains academic HS (age 31)	0.02	0.01	0.01	0.01	0.03	0.03
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains short college (age 31)	0.03**	0.02	0.02	0.02	0.03*	0.02
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains college (age 31)	0.01	0.01	0.01	0.01	0.02	0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Gross income (27-31 yrs old)	65.35**	54.11*	61.57	50.78	56.09	38.69
	(32.40)	(32.53)	(46.79)	(46.39)	(45.81)	(44.05)
Days/year unemp. (27-31 yrs old)	-1.45*	-1.31	-0.59	-0.74	-2.01	-1.49
	(0.86)	(0.88)	(1.04)	(1.05)	(1.46)	(1.52)
Observations	19622	19622	11845	11845	7777	7777
Schools	457	457	444	444	431	431

* p < 0.10, ** p < 0.05, *** p < 0.01

The table presents coefficient estimates and standard errors from the regression (1) of each outcome (in rows) on an indicator for focal child being surveyed interacted with sibling being in the critical grades after compulsory schooling, $g \in \{10, 11, 12\}$, controlling for municipality fixed effects, sibling grade, school size, and the indicator for focal child being surveyed interacted with siblings being in earlier and later grades, $g \in \{7, 8, 9\}$ and $g \in \{13, 14, 15, 16\}$, respectively. Additional controls include family background and municipality of birth. The main outcome is the standardized index for long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

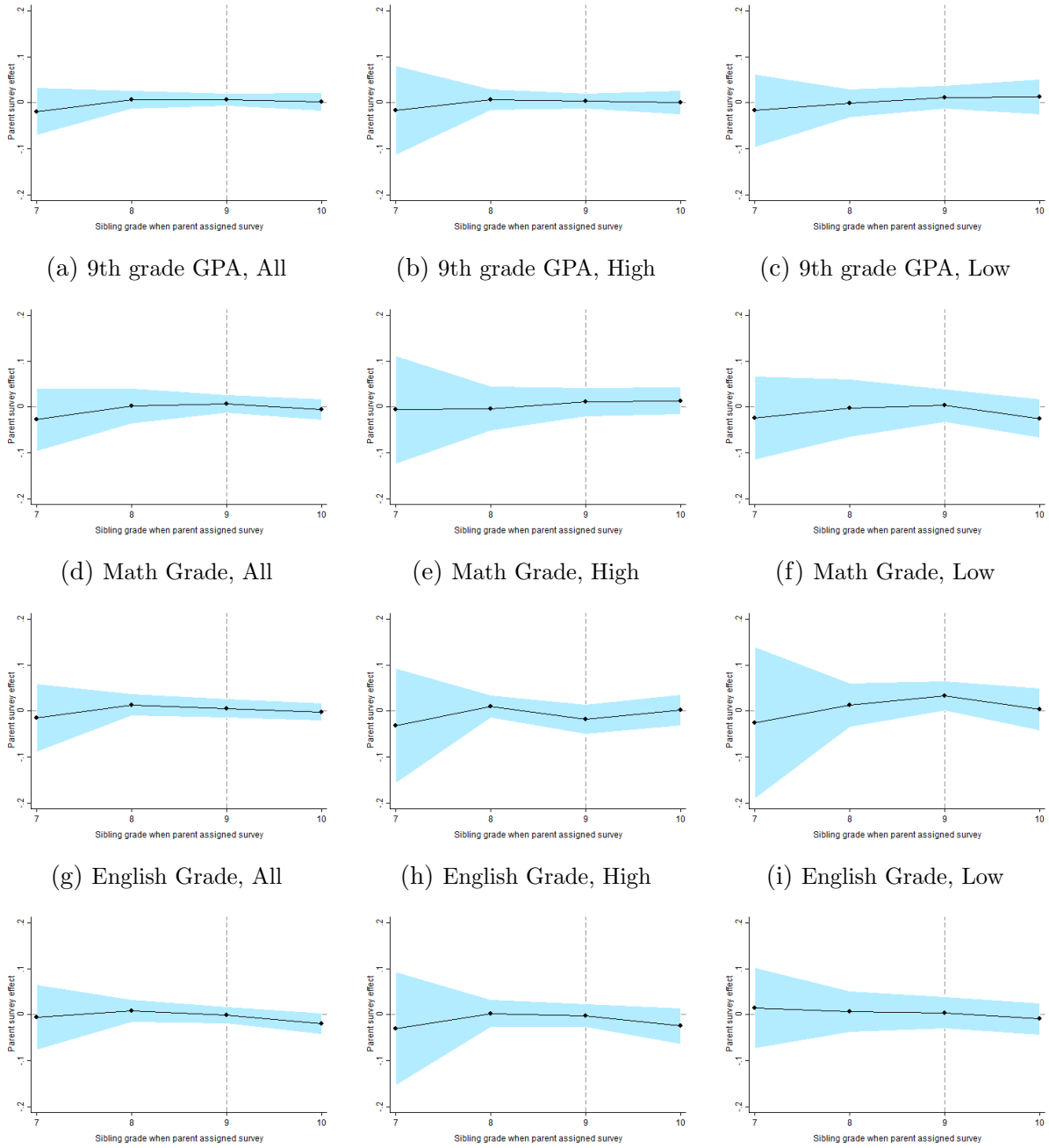
Table G.4: Effect of sibling at sampled school:
All Siblings of 1977 and 1982 cohort

Sample	All		High Education		Low Education	
Controls	No	Yes	No	Yes	No	Yes
Long-term index (std munic)	0.03	0.03*	0.04	0.05**	0.00	0.00
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
Attains high school (age 31)	0.01	0.01	0.02**	0.02*	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Attains academic HS (age 31)	0.00	-0.01	0.01	-0.01	-0.00	0.00
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains short college (age 31)	0.02	0.02	0.02	0.01	0.03	0.03
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Attains college (age 31)	0.02	0.02*	0.03	0.02	0.02	0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Gross income (27-31 yrs old)	23.05	43.98	27.23	67.32	-6.79	2.46
	(37.96)	(34.17)	(53.86)	(48.35)	(51.04)	(48.45)
Days/year unemp. (27-31 yrs old)	-1.45	-1.46	-2.51**	-3.10**	1.56	1.75
	(1.03)	(1.01)	(1.25)	(1.24)	(1.73)	(1.70)
Observations	19622	19622	11845	11845	7777	7777
Schools	457	457	444	444	431	431

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table presents coefficient estimates and standard errors from the regression (2) of each outcome (in rows) on an indicator for focal child being in a sampled school interacted with sibling being in the critical grades after compulsory schooling, $g \in \{10, 11, 12\}$, controlling for municipality fixed effects, sibling grade, school size, and the indicator for focal child being surveyed interacted with siblings being in earlier and later grades, $g \in \{7, 8, 9\}$ and $g \in \{13, 14, 15, 16\}$, respectively. Additional controls include family background and municipality of birth. The main outcome is the standardized index for long-term educational attainment and labor market performance. Effects on individual outcomes are presented below the index row. Income is expressed in 100 SEK, and prices are adjusted to 2014: SEK 1 = EUR 0.11 = USD 0.13. Standard errors clustered at the school level.

Figure G.2: Grades and GPA in 9th Grade,
Within-municipality - 1977 Cohort.



(a) 9th grade GPA, All (b) 9th grade GPA, High (c) 9th grade GPA, Low

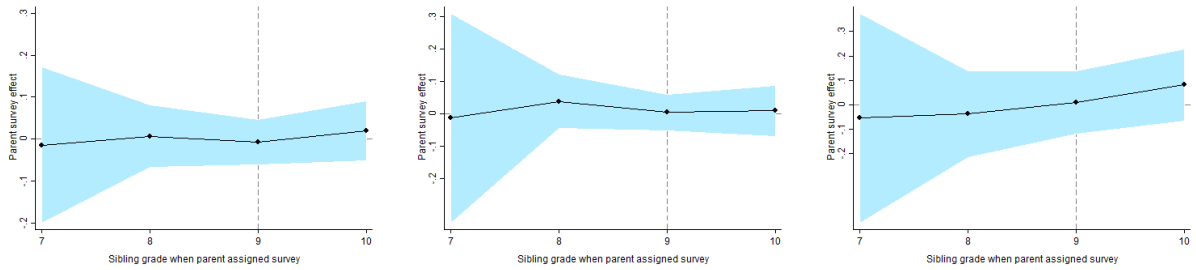
(d) Math Grade, All (e) Math Grade, High (f) Math Grade, Low

(g) English Grade, All (h) English Grade, High (i) English Grade, Low

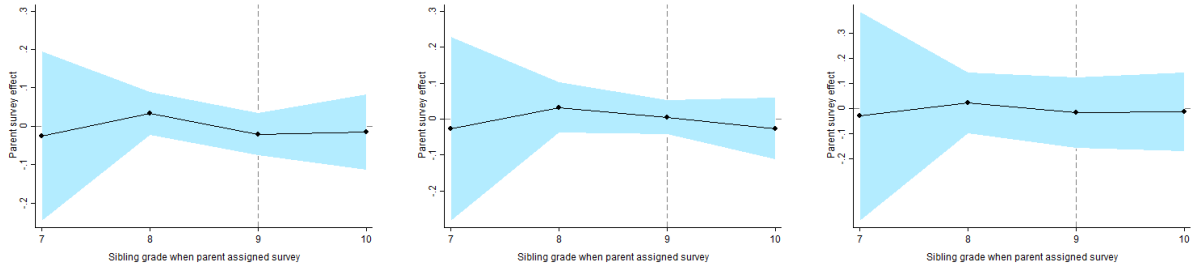
(j) Swedish Grade, All (k) Swedish Grade, High (l) Swedish Grade, Low

The Figure displays the parental survey effect on 9th grade performance measures of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: ETF77 cohort, Swedish born.

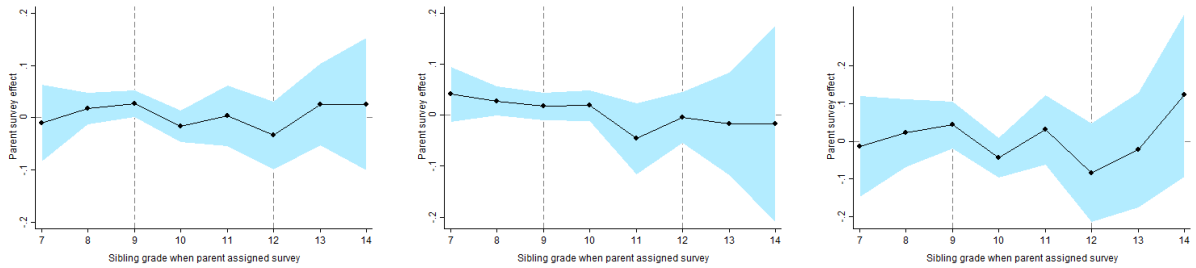
Figure G.3: 9th Grade Course Choices and High School Attainment, Within-municipality - 1977 Cohort



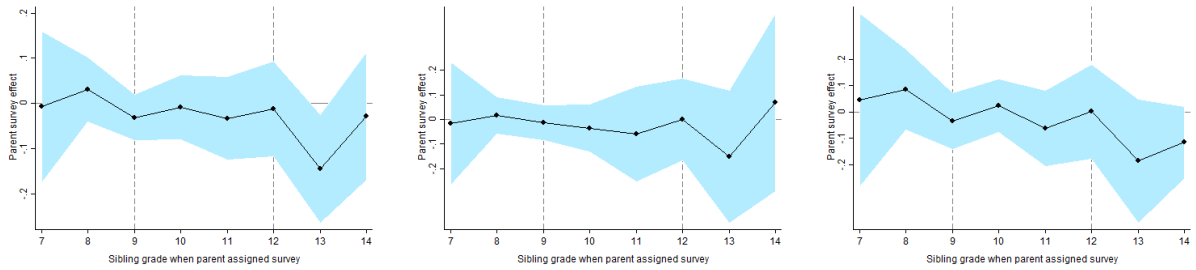
(a) Advanced Math, All. (b) Advanced Math, High. (c) Advanced Math, Low.



(d) Advanced English, All. (e) Advanced English, High. (f) Advanced English, Low.



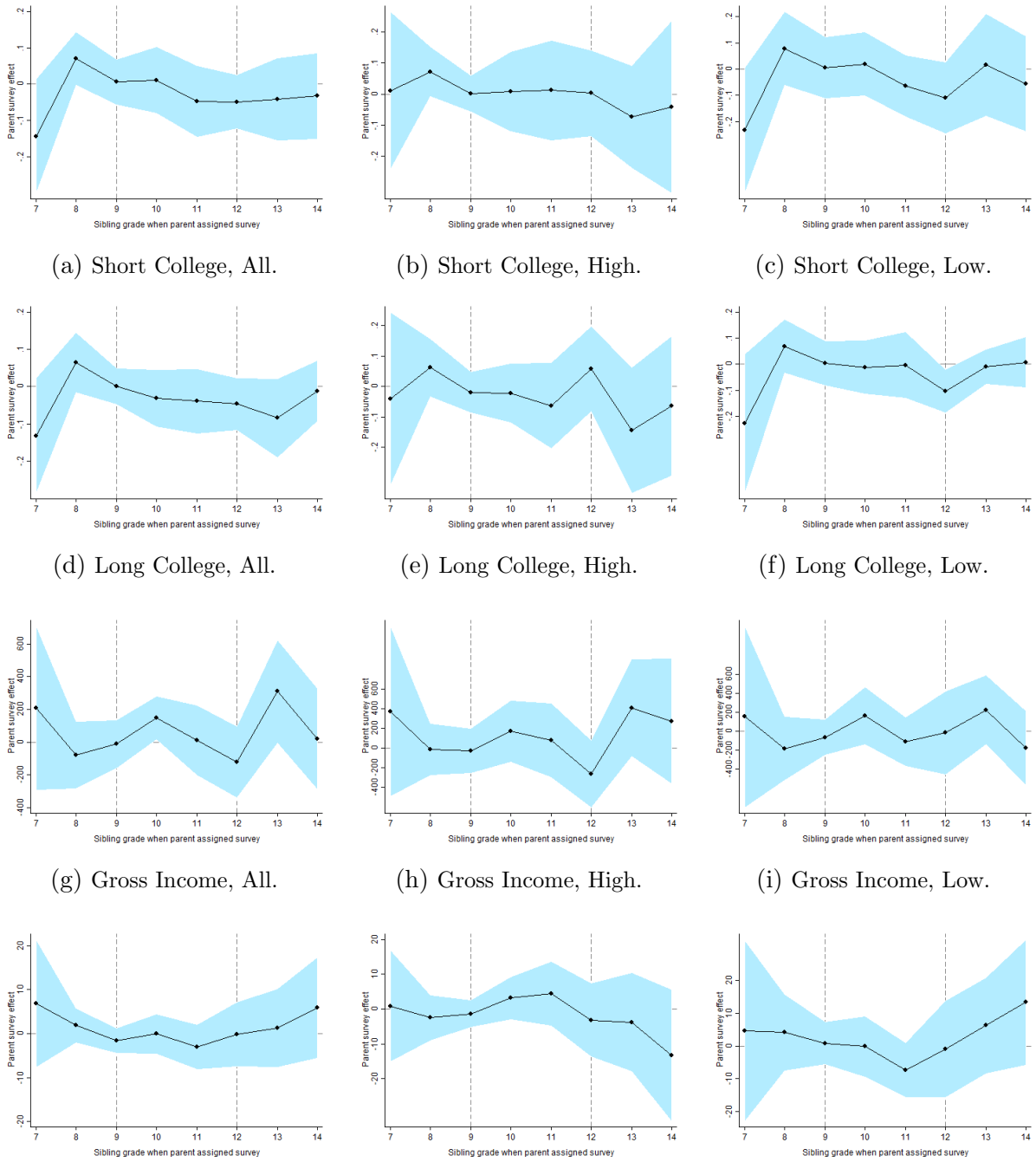
(g) High School, All. (h) High School, High. (i) High School, Low.



(j) Academic High School, All. (k) Academic High School, High. (l) Academic High School, Low.

The Figure displays the parental survey effect on 9th grade course choices and high school attainment by age 31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: ETF77 cohort, Swedish born.

Figure G.4: College Attainment and Labor Market Outcomes, Within-municipality - 1977 Cohort



(a) Short College, All. (b) Short College, High. (c) Short College, Low.

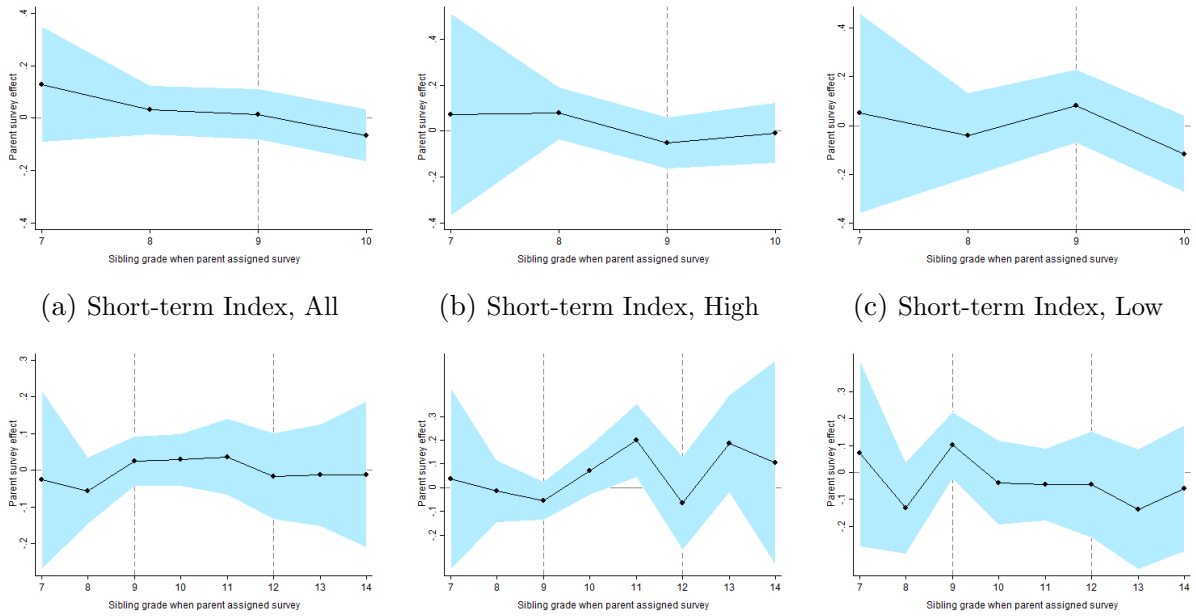
(d) Long College, All. (e) Long College, High. (f) Long College, Low.

(g) Gross Income, All. (h) Gross Income, High. (i) Gross Income, Low.

(j) Days Unemployed, All. (k) Days Unemployed, High. (l) Days Unemployed, Low.

The Figure displays the parental survey effect on college attainment by age 31, average yearly gross income at age 27-31, and average days unemployed during the year at age 27-31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: ETF77 cohort, Swedish born.

Figure G.5: Short- and Long-Term Effect of Parent Survey on Siblings, Between-school (indicator) - 1977 Cohort.

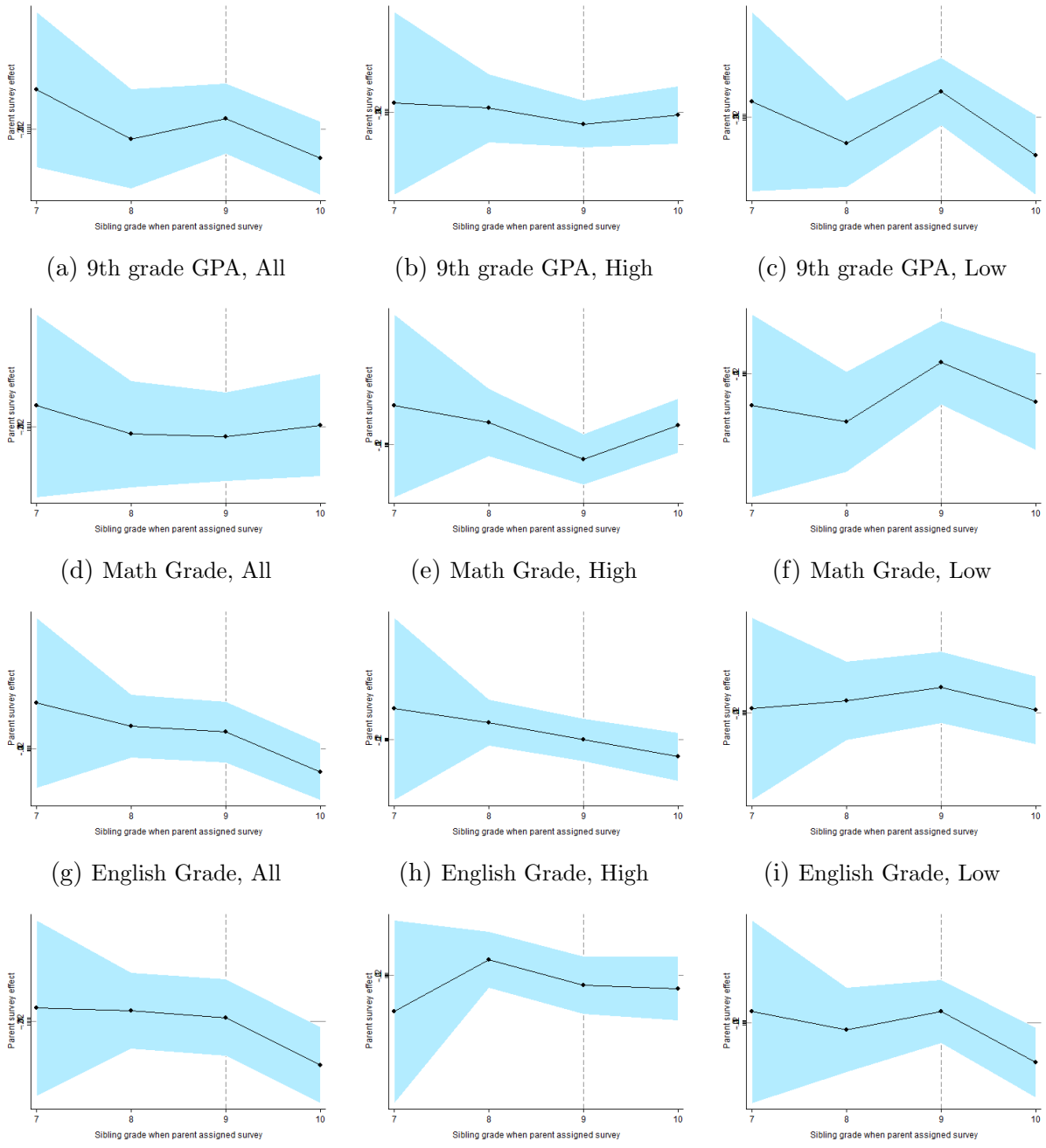


(a) Short-term Index, All (b) Short-term Index, High (c) Short-term Index, Low

(d) Long-term Index, All (e) Long-term Index, High (f) Long-term Index, Low

The Figure displays the parental survey effect on the short- and long-term index of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include school-cohort size and the full set of family background controls. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with $ETFSchool$ indicator for focal child being in a sampled school. Sample: ETF77 cohort.

Figure G.6: Grades and GPA in 9th Grade,
Between-school (indicator) - 1977 Cohort.



(a) 9th grade GPA, All (b) 9th grade GPA, High (c) 9th grade GPA, Low

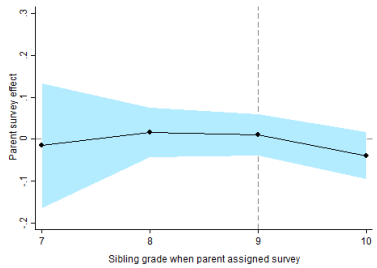
(d) Math Grade, All (e) Math Grade, High (f) Math Grade, Low

(g) English Grade, All (h) English Grade, High (i) English Grade, Low

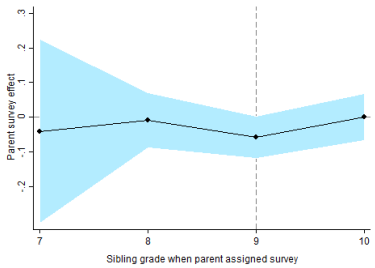
(j) Swedish Grade, All (k) Swedish Grade, High (l) Swedish Grade, Low

The Figure displays the parental survey effect on 9th grade performance measures of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with $ETFSchool$ indicator for focal child being in a sampled school. Sample: ETF77 cohort.

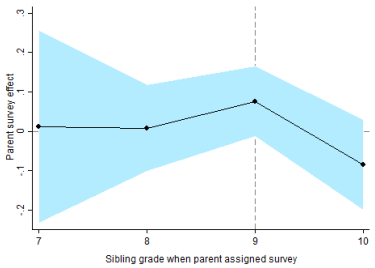
Figure G.7: 9th Grade Course Choices and High School Attainment, Between-school (indicator) - 1977 Cohort



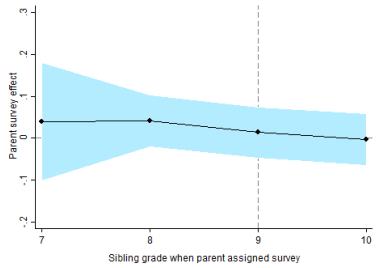
(a) Advanced Math, All.



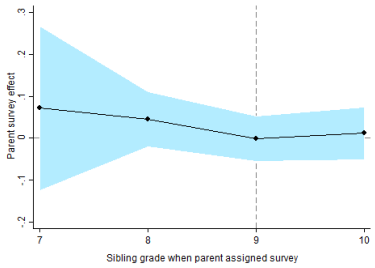
(b) Advanced Math, High.



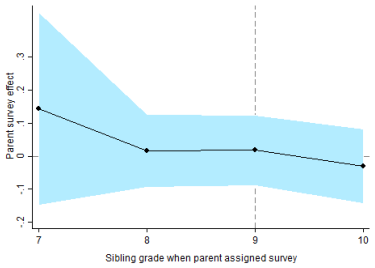
(c) Advanced Math, Low.



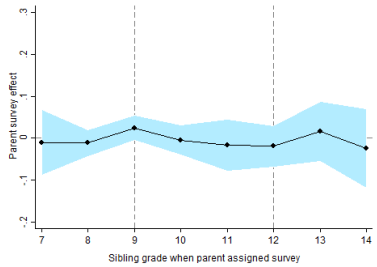
(d) Advanced English, All.



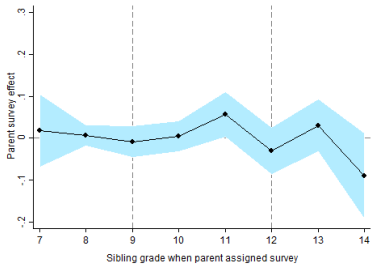
(e) Advanced English, High.



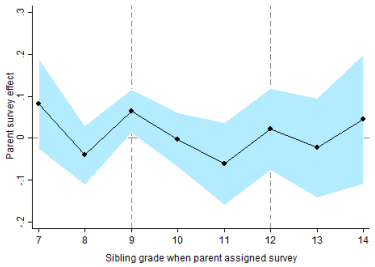
(f) Advanced English, Low.



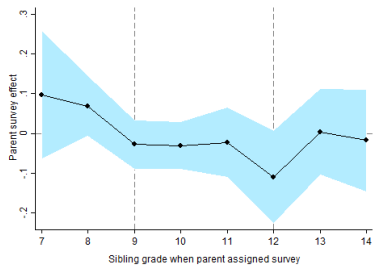
(g) High School, All.



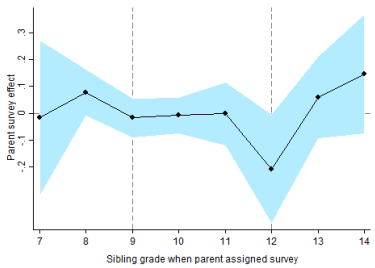
(h) High School, High.



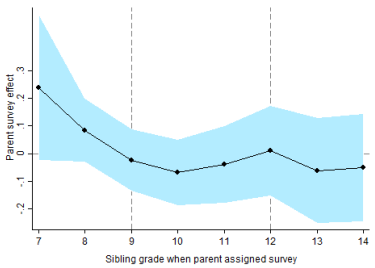
(i) High School, Low.



(j) Academic High School, All.



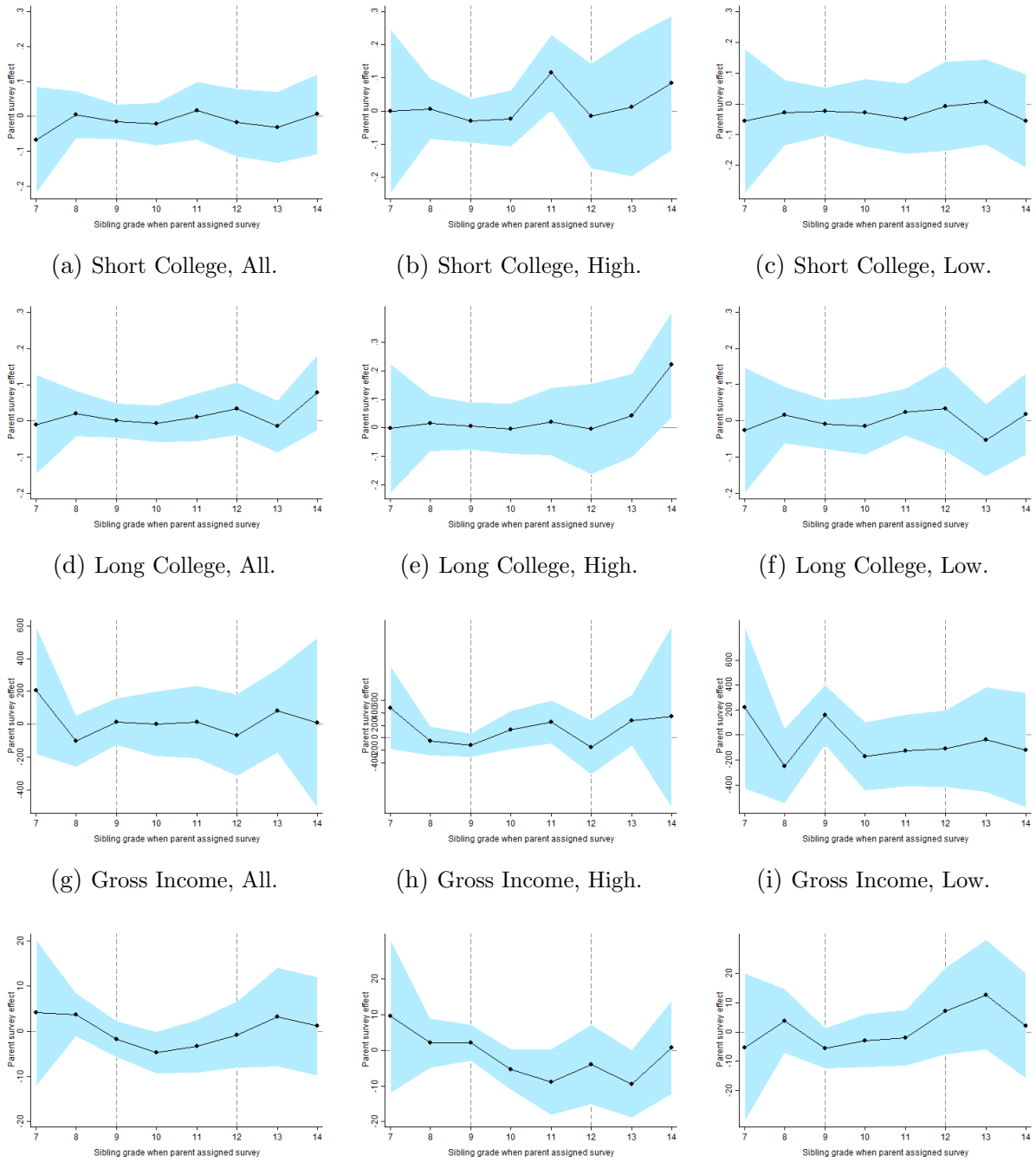
(k) Academic High School, High.



(l) Academic High School, Low.

The Figure displays the parental survey effect on 9th grade course choices and high school attainment by age 31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include school-cohort size and the full set of family background controls. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with *ETFSchool* indicator for focal child being in a sampled school. Sample: ETF77 cohort.

Figure G.8: College Attainment and Labor Market Outcomes,
Between-school (indicator) - 1977 Cohort



(a) Short College, All. (b) Short College, High. (c) Short College, Low.

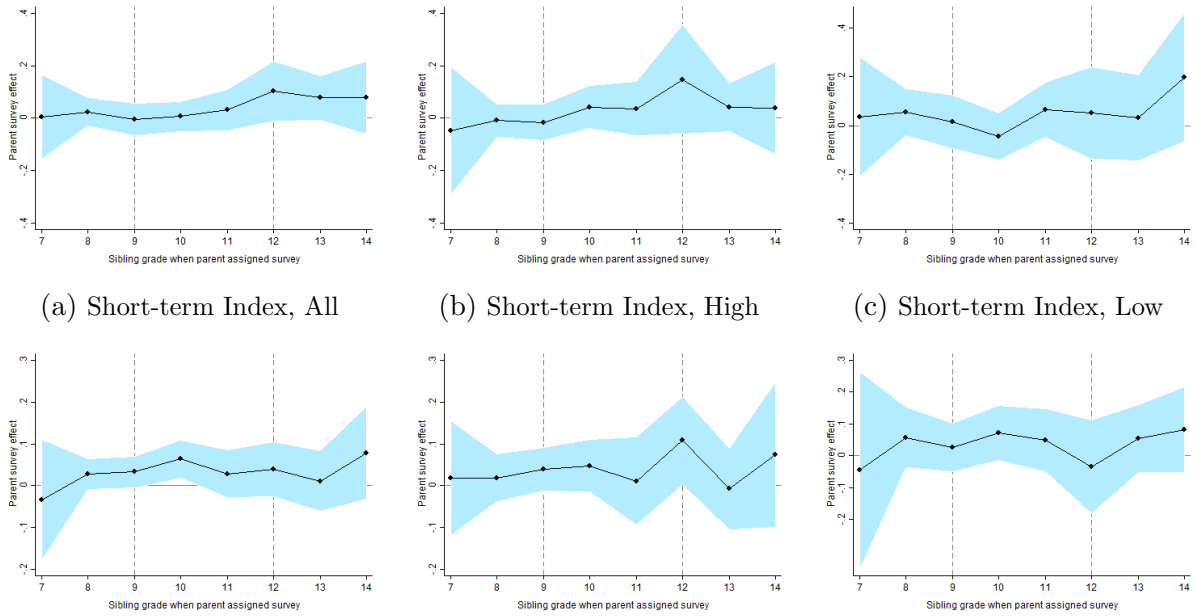
(d) Long College, All. (e) Long College, High. (f) Long College, Low.

(g) Gross Income, All. (h) Gross Income, High. (i) Gross Income, Low.

(j) Days Unemployed, All. (k) Days Unemployed, High. (l) Days Unemployed, Low.

The Figure displays the parental survey effect on college attainment by age 31, average yearly gross income at age 27-31, and average days unemployed during the year at age 27-31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with $ETFSchool$ indicator for focal child being in a sampled school. Sample: ETF77 cohort.

Figure G.9: Short- and Long-Term Effect of Parent Survey on Siblings, Within-Municipality - All.

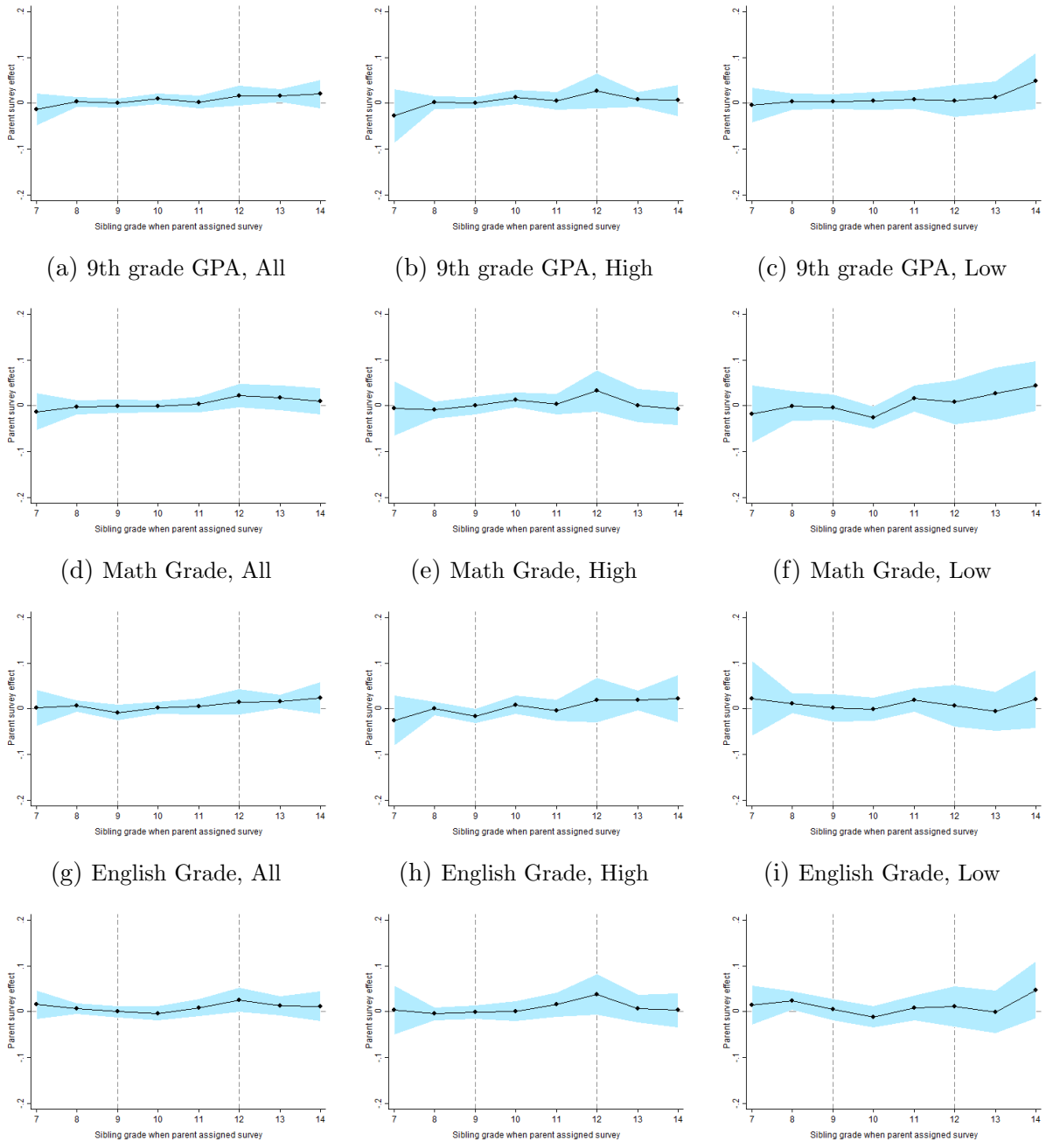


(a) Short-term Index, All (b) Short-term Index, High (c) Short-term Index, Low

(d) Long-term Index, All (e) Long-term Index, High (f) Long-term Index, Low

The Figure displays the parental survey effect on the short- and long-term index of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: Swedish born.

Figure G.10: Grades and GPA in 9th Grade,
Within-municipality - All.



(a) 9th grade GPA, All (b) 9th grade GPA, High (c) 9th grade GPA, Low

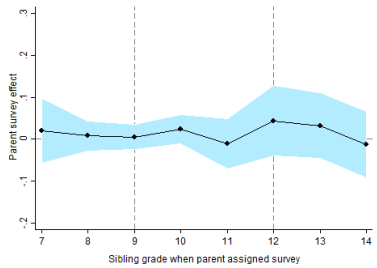
(d) Math Grade, All (e) Math Grade, High (f) Math Grade, Low

(g) English Grade, All (h) English Grade, High (i) English Grade, Low

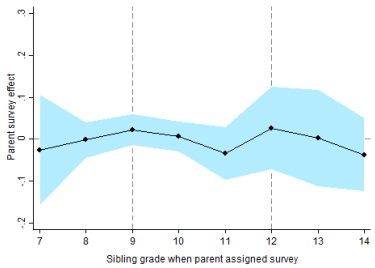
(j) Swedish Grade, All (k) Swedish Grade, High (l) Swedish Grade, Low

The Figure displays the parental survey effect on 9th grade performance measures of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: Swedish born.

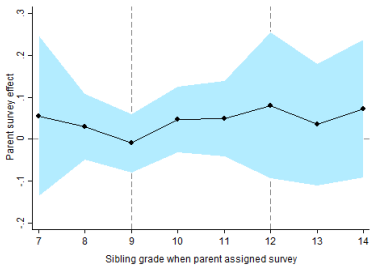
Figure G.11: 9th Grade Course Choices and High School Attainment, Within-municipality - All



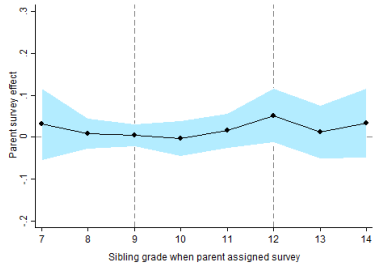
(a) Advanced Math, All.



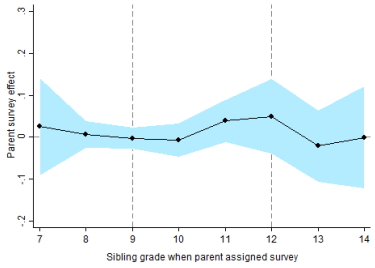
(b) Advanced Math, High.



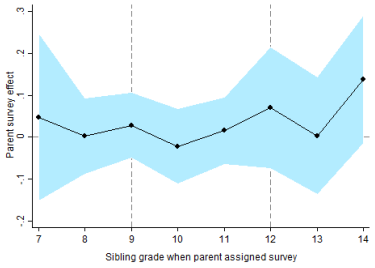
(c) Advanced Math, Low.



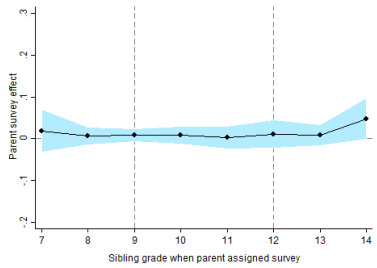
(d) Advanced English, All.



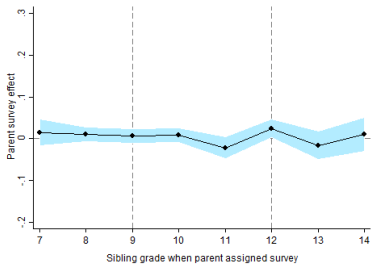
(e) Advanced English, High.



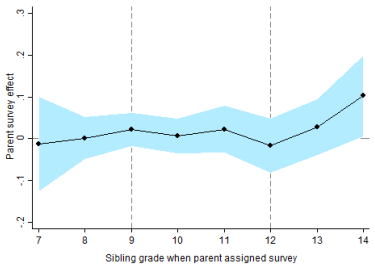
(f) Advanced English, Low.



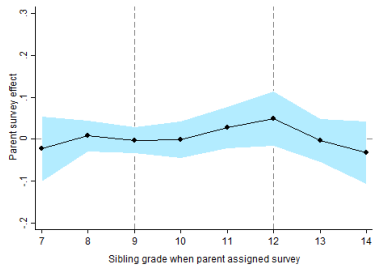
(g) High School, All.



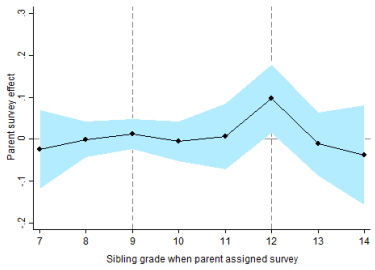
(h) High School, High.



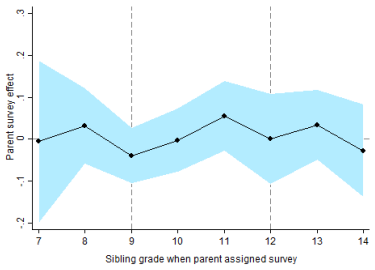
(i) High School, Low.



(j) Academic High School, All.



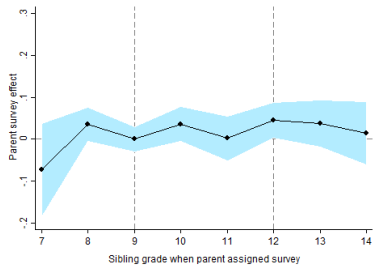
(k) Academic High School, High.



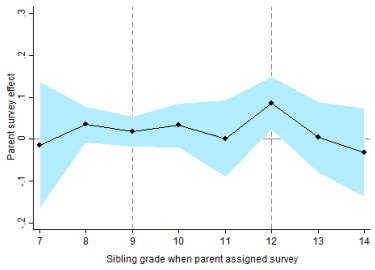
(l) Academic High School, Low.

The Figure displays the parental survey effect on 9th grade course choices and high school attainment by age 31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: Swedish born.

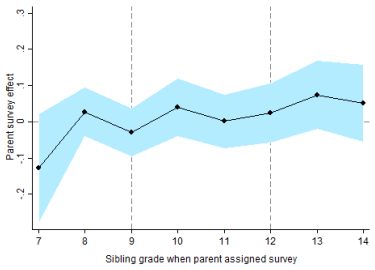
Figure G.12: College Attainment and Labor Market Outcomes, Within-municipality - All



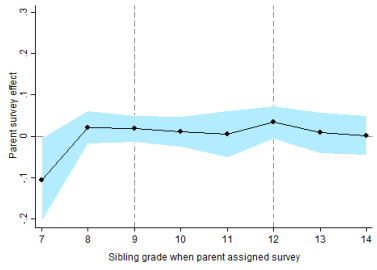
(a) Short College, All.



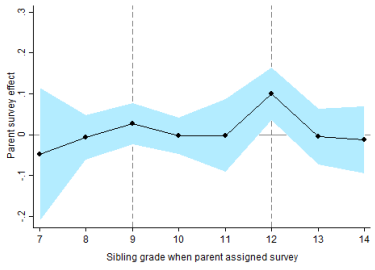
(b) Short College, High.



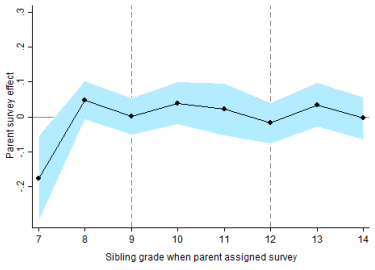
(c) Short College, Low.



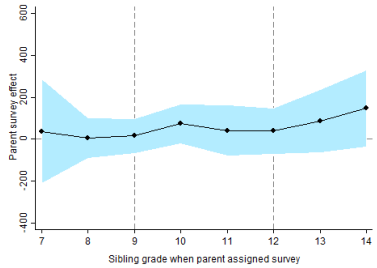
(d) Long College, All.



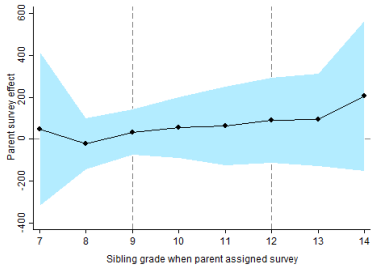
(e) Long College, High.



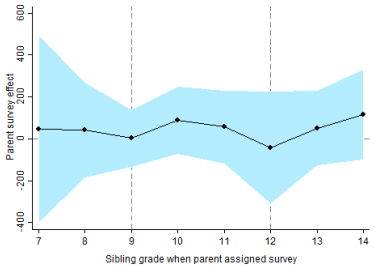
(f) Long College, Low.



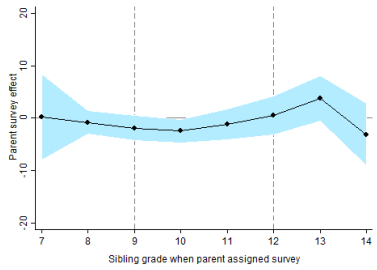
(g) Gross Income, All.



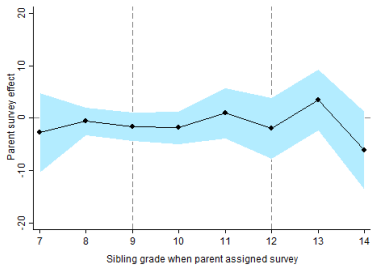
(h) Gross Income, High.



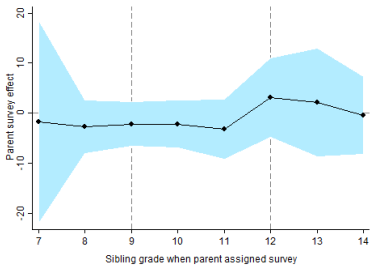
(i) Gross Income, Low.



(j) Days Unemployed, All.



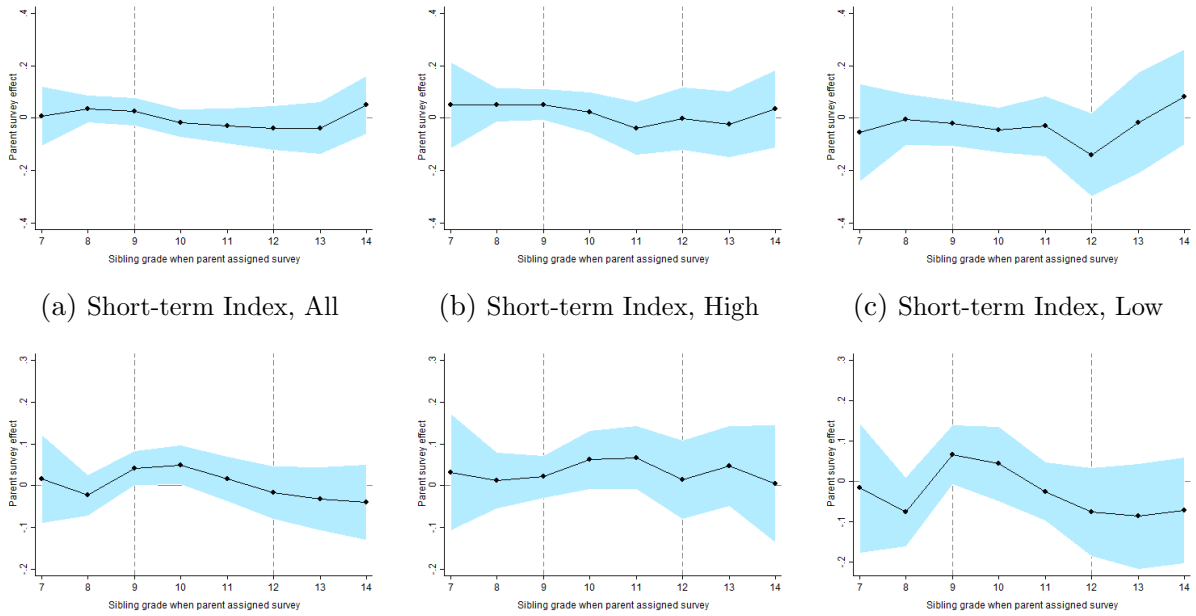
(k) Days Unemployed, High.



(l) Days Unemployed, Low.

The Figure displays the parental survey effect on college attainment by age 31, average yearly gross income at age 27-31, and average days unemployed during the year at age 27-31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Within-municipality, equation (1) with *Surveyed* indicator for focal child being surveyed. Sample: Swedish born.

Figure G.13: Short- and Long-Term Effect of Parent Survey on Siblings, Between-school (indicator) - All.

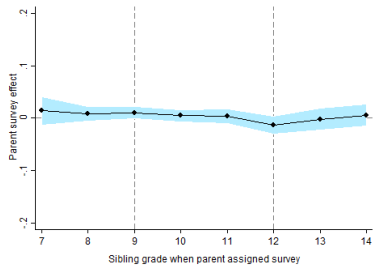


(a) Short-term Index, All (b) Short-term Index, High (c) Short-term Index, Low

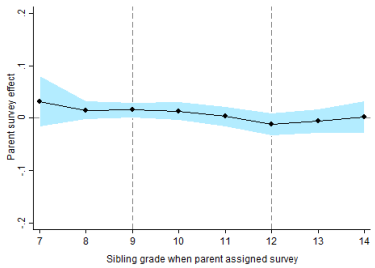
(d) Long-term Index, All (e) Long-term Index, High (f) Long-term Index, Low

The Figure displays the parental survey effect on the short- and long-term index of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include school-cohort size and the full set of family background controls. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with $ETFSchool$ indicator for focal child being in a sampled school. Sample: Full sample.

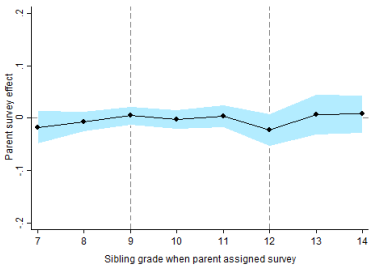
Figure G.14: Grades and GPA in 9th Grade,
Between-school (indicator) - All.



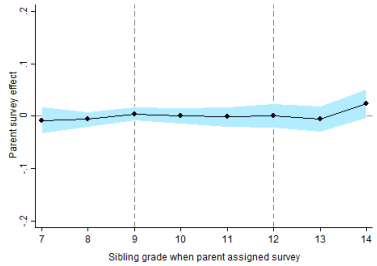
(a) 9th grade GPA, All



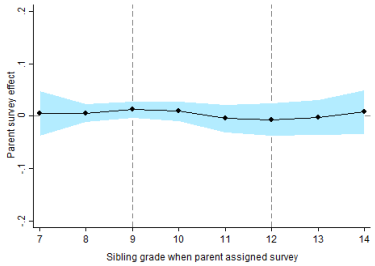
(b) 9th grade GPA, High



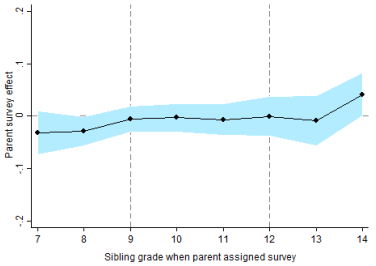
(c) 9th grade GPA, Low



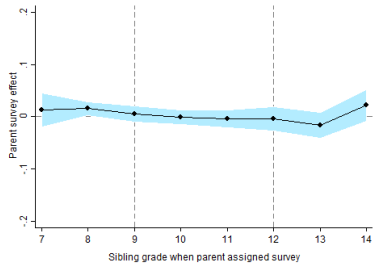
(d) Math Grade, All



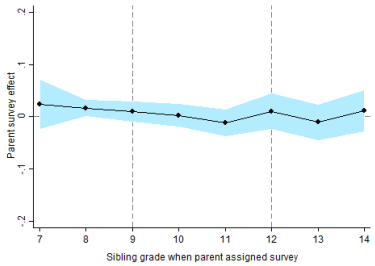
(e) Math Grade, High



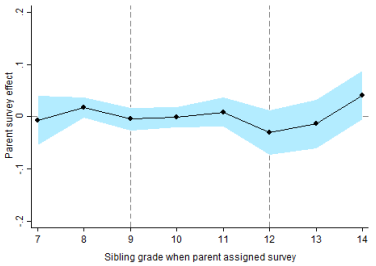
(f) Math Grade, Low



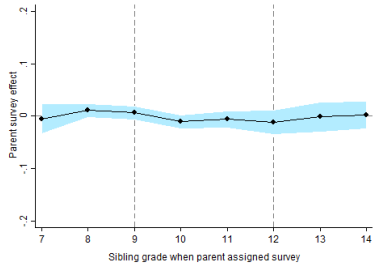
(g) English Grade, All



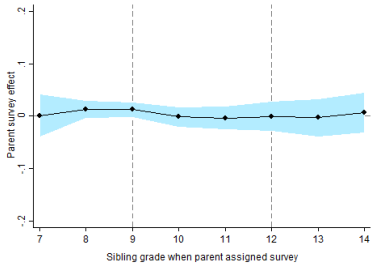
(h) English Grade, High



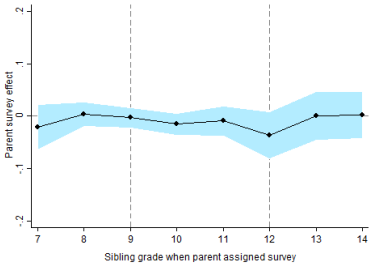
(i) English Grade, Low



(j) Swedish Grade, All



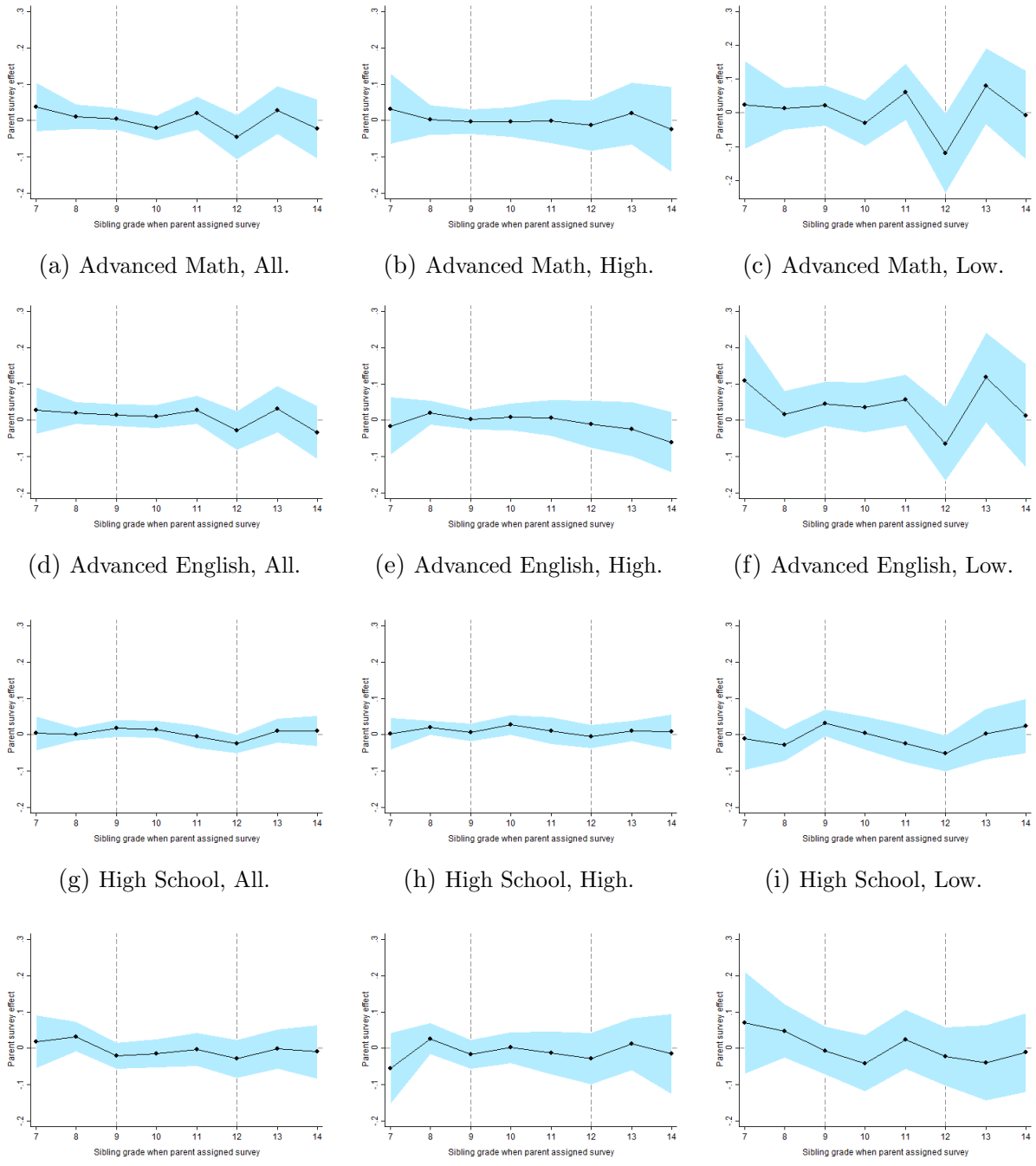
(k) Swedish Grade, High



(l) Swedish Grade, Low

The Figure displays the parental survey effect on 9th grade performance measures of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with $ETFSchool$ indicator for focal child being in a sampled school. Sample: Full sample.

Figure G.15: 9th Grade Course Choices and High School Attainment, Between-school (indicator) - All



(a) Advanced Math, All. (b) Advanced Math, High. (c) Advanced Math, Low.

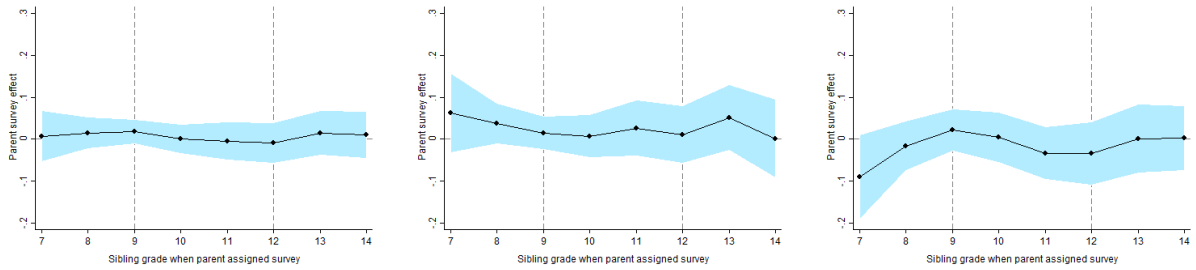
(d) Advanced English, All. (e) Advanced English, High. (f) Advanced English, Low.

(g) High School, All. (h) High School, High. (i) High School, Low.

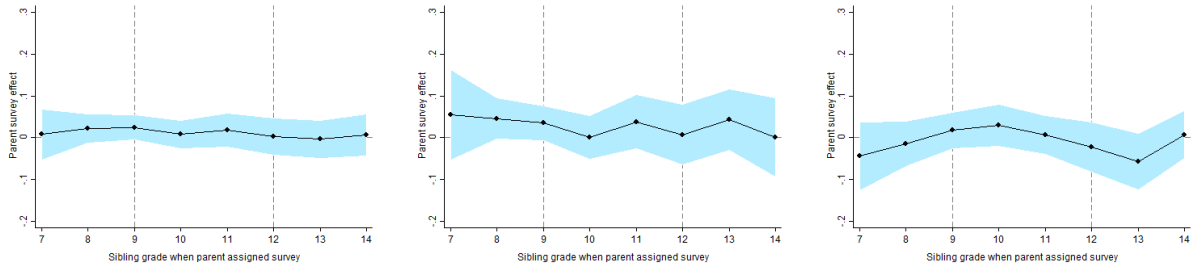
(j) Academic High School, All. (k) Academic High School, High. (l) Academic High School, Low.

The Figure displays the parental survey effect on 9th grade course choices and high school attainment by age 31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include school-cohort size and the full set of family background controls. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with *ETFSchool* indicator for focal child being in a sampled school. Sample: Full sample.

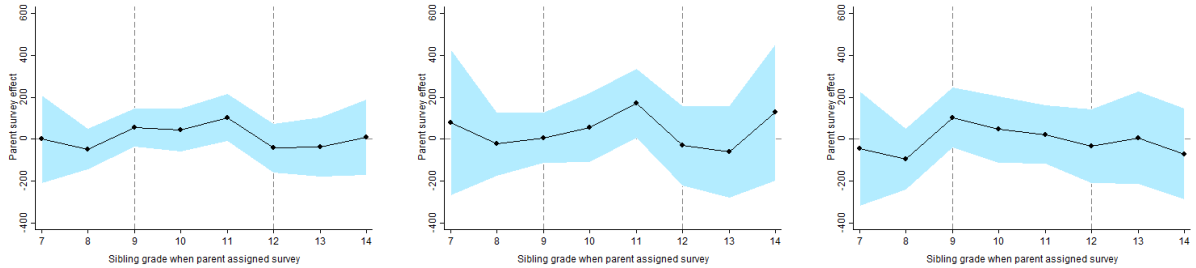
Figure G.16: College Attainment and Labor Market Outcomes,
Between-school (indicator) - All



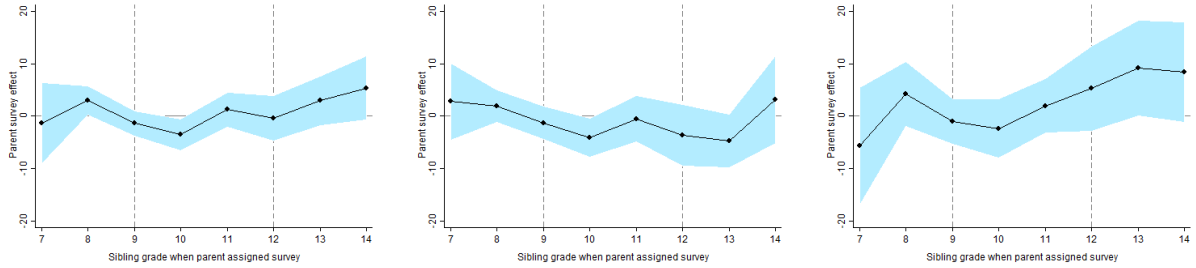
(a) Short College, All. (b) Short College, High. (c) Short College, Low.



(d) Long College, All. (e) Long College, High. (f) Long College, Low.



(g) Gross Income, All. (h) Gross Income, High. (i) Gross Income, Low.



(j) Days Unemployed, All. (k) Days Unemployed, High. (l) Days Unemployed, Low.

The Figure displays the parental survey effect on college attainment by age 31, average yearly gross income at age 27-31, and average days unemployed during the year at age 27-31 of the focal child's sibling (y-axis) by stipulated grade of the sibling when their parent was surveyed (x-axis). The vertical dashed lines mark 9th and 12th grade; i.e. the last year of compulsory schooling and high school, respectively. Regressions include municipality fixed effects, school-cohort size, and the full set of family background controls and municipality at birth fixed effects. The first column displays the figures for the full sample, the second column for parents with high education, and the third column for parents with low education. Specification: Between-school, equation (2) with $ETFSchool$ indicator for focal child being in a sampled school. Sample: Full sample.