WORKING PAPER 6/2026

Academic or Vocational Education? Parents' Expectation and Preference at Age 13 and Their Children's Adult Educational Attainment and Earnings in a Swedish School Year Cohort

Daniela Andrén

ISSN 1403-0586

Örebro University School of Business
SE-701 82 Örebro, Sweden

Academic or Vocational Education? Parents' Expectation and Preference at Age 13 and Their Children's Adult Educational Attainment and Earnings in a Swedish School Year Cohort

Daniela Andrén

Örebro University, Sweden

daniela.andren@oru.se

Abstract

When their children were 13, the parents of a school-year cohort in a midsize town in middle Sweden reported which schooling orientation (academic or vocational) they *preferred* for the child and which they *expected* the child would take. This paper interprets these two answers as proxies for parental investment in schooling: preference captures the direction of the aim, while expectation captures how far the parent believes it will be realized. The research design cannot disentangle parental *push* from parental *foresight*, so the estimates are interpreted as conditional associations. Instead, it leverages the joint measurement of preferences and expectations at the same age, rich background controls, and long-run register follow-up to describe how these two dimensions of parental schooling orientation relate to their children's later attainment and labor-market outcomes. Linking the cohort to Swedish administrative registers half a century later and conditioning on socioeconomic group, parental education, sex, birth-year group, and measured cognition at the same age, parental *expectation* is associated with higher adult schooling, while parental *preference*, holding expectation fixed, is more closely related to mid-career earnings than to schooling, with no corresponding association with late-life subjective well-being. Parents' post-compulsory expectation at age 16 maps onto their children's realized schooling in a steep, monotone gradient, from a 7 percent post-secondary share where parents expected immediate work to 75 percent where they expected a three-year academic upper-secondary program.

Keywords: parental expectations; parental preferences; educational attainment; earnings.

JEL: I21; I24; J24; J31.

1 Introduction

In the autumn of 1968, the parents of a school-year cohort in a midsize town in middle Sweden filled in a questionnaire about their 13-year-old children. Two of the items asked which schooling orientation, academic (gymnasium-preparatory) or vocational, the parents *preferred* for the child, and which one they *expected* the child would actually take. Roughly 70% parents preferred the academic orientation and roughly 60% expected it. This paper reads the two answers not as forecasts to be scored for accuracy but as *proxies for parental investment*: how hard, and in which direction, a family pushes a child toward a schooling orientation, through the encouragement, monitoring, and resources that a parent who wants and expects the academic route brings to bear, and that a parent who wants neither does not.

A large literature in the economics of education (Cunha and Heckman, 2007, 2008; Heckman et al., 2006) treats parental investment as a central input into children's human-capital formation, and parental expectations and preferences as observable expressions of it. Yet the two are rarely measured separately, in the same survey, at the same age, and on the same scale. When only one is observed, it is unclear whether a reported "academic future" is an expectation ("this is where my child is heading") or a preference ("this is where I would like my child to head"). The distinction matters: expectations aggregate parental private information about the child's ability, effort, and circumstances, while preferences may reflect values, status concerns, or ambitions that need not coincide with the child's trajectory (Schwandt, 2016). The Individual Development and Adaptation program asked both questions in the same questionnaire, making the two jointly observable. This paper asks whether these two items track realized adult attainment once family background and the child's measured cognition at the same age are held fixed. The age-13 responses are linked to outcomes measured up to 54 years later: schooling and earnings in Swedish administrative registers between 1990 and 2022, and subjective well-being measured by postal questionnaire when the cohort reached age 67.

The two measures relate to different adult margins. Conditional on the family's socioeconomic group, family education, an intact home at grade three, sex, birth-year group, and a same-wave measure of cognition, parental *expectation* is associated with higher adult schooling, 0.29 of a level on the six-level Swedish education scale and roughly 13 percentage points higher post-secondary attainment. Parental *preference*, holding expectation fixed, is only weakly associated with schooling but is associated with roughly 40 thousand Swedish kronor (SEK) higher mean annual earnings over ages 45–54; in contrast, expectation is unrelated to

earnings.[1] Adding post-exposure school grades at 16 attenuates the expectation–schooling association from 0.29 to 0.19 of a level and it is no longer statistically significant. Late-life well-being results are unchanged under inverse-probability-of-response weighting. A direct test does not reject equality of the expectation and preference coefficients, so the outcome-specific split is treated as a regularity in this setting, not an established difference. A parental preference without a matching expectation nonetheless relates to an earnings advantage decades later; a realized academic preference carries no penalty or premium in life satisfaction or loneliness late in life.

The contribution is threefold. First, the paper uses separately measured parental expectations and preferences (on the same children and at the same age) as proxies for parental investment in a schooling orientation, and shows that they track different adult outcomes. Second, it documents that the answers reflect a followed direction rather than idle hope: parents' concrete post-compulsory expectation at 16 maps onto the schooling their children actually reached in the registers in a steep, monotone gradient. Third, it asks whether an unmet parental preference leaves a subjective "scar" at retirement age, and finds none. Throughout, the design cannot separate parental *push* from parental *foresight*, so the estimates are interpreted as conditional associations consistent with an investment channel, not as causal effects.

The remainder of the paper is organized as follows. Section 2 sets out the conceptual framework and maps the model's objects to the measured variables. Section 3 describes the cohort, registers, variables, and institutional setting (the comprehensive-school reform, track structure, and labor-market institutions framing the 1990–2022 register window). Section 4 states the identification strategy. Section 5 reports the main results, including the pre-specified preference-versus-expectation contrast and the comparison of parental expectations with adult educational reality. Section 6 presents robustness and sensitivity analyses. Section 7 interprets the findings, and Section 8 concludes; the appendix collects supporting material.

2 Conceptual framework

2.1 The parental expectation and preference: two readings

The framework combines three standard building blocks: the human-capital investment model, in which schooling is chosen by comparing discounted benefits and costs; the Mincer earnings function, in which earnings increase in completed schooling; and the technology of

[1] As of July 2026, the conversion rates are 1 SEK $\approx$ 0.10 USD and 1 SEK $\approx$ 0.09 EUR.

skill formation, in which a child's adult skills are produced from early endowments and a stream of parental and institutional investments (Cunha and Heckman, 2007, 2008; Heckman et al., 2006). Within this apparatus, a parent asked about a child's schooling orientation (academic or vocational) can be read as reporting one of two distinct quantities.

A parental *expectation* ("which orientation will the child actually take?") aggregates the parent's private information about the child's ability, effort, and circumstances, much of which is unobserved by the analyst. To the extent the expectation is accurate, it should relate to the child's realized position in the educational system, namely the schooling the child will go on to complete. In this sense, the expectation is a proxy for the information set the educational system itself will act on when it selects and sorts children.

A parental *preference* ("which orientation do you want for the child?") is what the parent wants for the child. Conditional on the expectation, the residual variation in the preference reflects what the parent wants over and above what the parent expects. This residual may encode parental ambition, the family's valuation of academic attainment, or a push toward outcomes the child is not currently on course to reach (Schwandt, 2016; Sewell et al., 1969). Holding the expectation fixed, the preference need not relate to the same outcome the expectation relates to; but, as the model makes precise, anything that raises schooling must also raise earnings through schooling, so the two channels are connected, not exclusive.

These two readings share a common substrate: parental *investment*. A parent who wants the academic route and expects the child to reach it is more likely to invest toward it (encouraging, monitoring, discussing the choice, and directing time and attention into the child's schooling), while a parent who wants neither invests less in that direction. Read this way, the preference marks the *direction* of the investment and the expectation its *intensity* or perceived success. The investment reading and the information reading are not mutually exclusive, and the design cannot separate them: an expectation that relates to later schooling may do so because the parent *pushed* the child there (investment) or because the parent *read* an ability the school system would have sorted on anyway (foresight). Section 6 bears on the distinction (conditioning on the child's measured cognition leaves the associations standing, while additionally controlling for post-exposure school performance attenuates the expectation–schooling link), but cannot settle it. The paper therefore interprets its estimates as conditional associations consistent with a parental-investment channel, and is explicit about what it cannot rule out.

2.2 A compact formal framework

Let θ_i denote the child's stock of skills relevant to educational selection at the time the parental items are elicited, produced from earlier endowments and investments by the technology of skill formation (Cunha and Heckman, 2007, 2008). The parents observe a private signal of this stock,

$$s_i = \theta_i + \varepsilon_i, \tag{1}$$

which aggregates information the analyst does not see (effort, motivation, circumstances at home). The educational system selects children into successively higher levels of schooling on realized performance in the same skill stock. Write S_i for the completed level (the six-level scale of the registers), increasing in $\theta_i + u_i$, where u_i collects shocks realized after the items are collected, and write

$$C_i = \mathbf{1}\{S_i \geq \bar{s}\} \tag{2}$$

for the post-secondary margin. Adult earnings follow a Mincer-type equation,

$$Y_i = \mu + \rho\, S_i + \lambda\, N_i + \eta_i, \qquad \rho > 0,\ \lambda > 0, \tag{3}$$

increasing in completed schooling with return ρ (the central mechanism of the human-capital model) and in a bundle of labor-market-valued skills N_i (persistence, sociability, ambition) that the diploma does not certify (Heckman et al., 2006).

The parent's *expectation* is where they expect the selection process to place the child, given their information,

$$E_i = \mathbf{1}\Big\{\Pr(C_i = 1 \mid s_i, \mathbf{X}_i) \geq \pi\Big\}, \tag{4}$$

for some reporting threshold π. Up to reporting error, the expectation is a sufficient statistic for the parent's private information about the selection-relevant skill stock. The parent's *preference* is the human-capital investment comparison (the discounted benefits and costs of the academic orientation as the family sees them), shifted by a family valuation term,

$$P_i = \mathbf{1}\Big\{V(\text{academic} \mid s_i, \mathbf{X}_i) + \nu_i \geq V(\text{vocational} \mid s_i, \mathbf{X}_i)\Big\}, \tag{5}$$

where ν_i captures what the family wants over and above what it expects (ambition, the psychic value it attaches to academic attainment, status concerns). Conditional on the expectation and on a direct measure of the child's contemporaneous cognition, the residual variation in the preference is variation in ν_i. The valuation term matters for outcomes

because parental investments respond to it,

$$N_i = N\Big(I_i, \theta_i\Big), \qquad I_i \text{ increasing in } \nu_i. \tag{6}$$

a family that wants the academic orientation more than it expects it invests (encouragement, monitoring, aspiration-setting), and much of that investment builds the uncertified skills N_i that equation (3) prices (Cunha and Heckman, 2008; Heckman et al., 2006).

2.3 Hypotheses

The model, read together with the status-attainment evidence on parental encouragement (Sewell et al., 1969) and evidence that noncognitive skills carry labor-market returns beyond schooling (Heckman et al., 2006), yields three hypotheses about how parental expectation and preference relate to the child's adult outcomes. One fact shapes all of them: because earnings rise with completed schooling (equation 3), anything associated with more schooling is thereby associated with higher earnings. The model therefore does *not* restrict either measure to a single outcome; it restricts the *channel* each works through: expectation through the schooling the education system awards, preference additionally through skills the labor market rewards directly.

Hypothesis 1 (H1) concerns schooling. A parent who expects the child to take the academic orientation should be associated with the child completing more schooling; and of the two measures, expectation should be the stronger one for schooling, because it summarizes the parent's reading of the child's actual prospects (the same information the school system acts on when it admits and promotes students), whereas preference is what the family wants, which need not match where the child is actually heading.

Hypothesis 2 (H2) concerns earnings. Holding expectation and controls fixed, a parent's preference for the academic orientation should be associated with higher mid-career earnings; and of the two measures, preference should be the stronger one for earnings, because the part of preference not already captured by expectation reflects the family's ambition and valuation of achievement, which prompt investments (encouragement, monitoring, aspiration-setting) that build skills employers reward but no diploma records. A sharper implication follows: once the child's completed schooling is *also* held fixed, expectation's association with earnings should fall toward zero, because expectation reaches earnings only by way of the diploma, while a positive preference association may survive through uncertified skills.

Hypothesis 3 (H3) concerns interaction between preference and expectation. The associations between preference and expectation should be larger for families whose expectation and preference *agree* on the academic orientation than the two measures would produce

separately. In the language of skill formation this is dynamic complementarity (Cunha and Heckman, 2007, 2008): the extra investment a family makes when it wants the academic orientation pays off more when the child is also expected to reach it.

H1 and H2 are the paper's *outcome-specific split*: what parents expect and what parents prefer, though almost the same question, carry information about *different* adult outcomes (schooling for expectation, earnings for preference). The claim has one direct empirical form. Summarize the two measures by how far the family leans toward the academic orientation overall (do neither, one, or both of the expectation and the preference point to it?) and by the *gap* between them (does the family expect the academic orientation more than it prefers it, or prefer it more than it expects it?). The outcome-specific split then implies that the gap works in *opposite* directions for the two outcomes: it tilts toward expectation where expectation leads (schooling) and toward preference where preference leads (earnings). A single gap that changes sign between schooling and earnings is the sharpest statement of the claim. H3 is a secondary implication of the same investment technology. Section 4 states the regressions (and defines the coefficients) that test each hypothesis.

A secondary, exploratory question concerns the affective consequences of an unmet preference. The unmet-preference account of well-being holds that the gap between what one hoped for and what one attained can depress subjective well-being (Schwandt, 2016; Clark et al., 2008; Layard et al., 2014). Translated to the parent generation's hopes for the child, this reading implies that a child whose parents preferred an academic future that did not materialize might report lower late-life well-being than a child whose parental academic preference was realized. This implication is treated as exploratory (it is not derived from the model above) and is reported with the multiplicity caution described in Section 4.

3 Data and institutional setting

3.1 The IDA cohort and data sources

The data come from the Individual Development and Adaptation (IDA) research program, a longitudinal study that followed a school-year cohort in a midsize town in middle Sweden (Magnusson, 1988; Bergman et al., 2018). The cohort is defined by school year, not by birth year: it comprises all children who were in the third grade in the spring of 1965, together with classmates who joined the cohort's grade-level classes through the lower-secondary years and were integrated as full program members. Cumulative membership is 1,392 individuals (710 men and 682 women). Most members were born in 1955, but a minority were born earlier or later, so birth year enters the analysis only as a three-level grouping (Section 3.3). All

cumulative members are followed in the national registers; the register-eligible universe for outcomes drawn from the Longitudinal Integrated Database for Social Insurance and Labour Market Studies (LISA), which begins in 1990, is 1,348 individuals.

The analysis of this paper combines three sources. The two age-13 parent-reported items and supporting age-13 and age-16 items come from the parent and pupil questionnaires of the program's lower-secondary waves (collected in 1968 at age 13 and around 1971 at age 16). Educational attainment, earnings, and employment come from the LISA register, annually from 1990 to 2022. Subjective well-being is measured in a retirement-age postal questionnaire administered when the cohort was 67 (in 2022–23). Construction of the analytic file from these sources is documented in Appendix B.

The cohort's municipality of residence is not named in this paper, although the IDA program and prior publications using these data routinely name it. Because most cohort members are alive, naming the municipality would invite geographic re-identification; the paper therefore describes the setting only as "a midsize town in middle Sweden." This is a stricter standard than the rest of the IDA program applies and is adopted here to protect respondents.

3.2 Institutional setting

The comprehensive-school reform. The cohort passed through the Swedish school system while that system was being rebuilt. The 1949–1962 compulsory-school reform consolidated the country's seven-year *folkskola* and the parallel *realskola*–gymnasium streams into a single nine-year comprehensive school (*grundskola*), extending compulsory schooling from seven to nine years; the staged municipal rollout of the reform has been the subject of a series of long-run-effect studies (Meghir and Palme, 2005; Holmlund, 2007; Lundborg et al., 2014). The cohort's town completed the transition to the unified *grundskola* during the 1960s and early 1970s, so the cohort (in grade three in spring 1965, grade six in spring 1968 when the two parental items were recorded, and grade nine in 1971) belongs to the first generation schooled essentially entirely inside the comprehensive system.

The track choice at the end of compulsory school. Under the 1962 comprehensive-school curriculum in force when the parental items were collected, differentiation entered in the upper grades: pupils chose elective courses in grades seven and eight, and grade nine was organized in distinct lines. The academically oriented lines prepared for the gymnasium, while general and practically oriented lines led toward vocational routes and work. In the Swedish terminology of the day, the 1968 parent items asked about the "theoretical" versus the "practical" line. To avoid any confusion with the theoretical model of Section 2, this

paper calls the theoretical, gymnasium-preparatory route the *academic orientation* and the practical, vocational route the *vocational orientation*, and refers to the two survey items as the parent's academic *preference* and academic *expectation*.

Two features of the moment matter for reading the parent's answers. First, the choice was consequential: the grade-nine line governed, in practice, access to long gymnasium programs. Second, the structure itself was in motion: the curriculum revision adopted in 1969 began dismantling the grade-nine lines in favor of free elective choice in the years immediately after the items were collected. Parents were therefore assessing a track choice that was real but whose institutional form was being softened as their children approached it, one more reason to read the two items as parental information and preference rather than as mechanical anticipations of an administrative assignment.

Post-compulsory education. The post-compulsory menu facing the cohort at the end of grade nine (1971) consisted of the long, university-preparing gymnasium lines of three to four years; the two-year continuation school (*fackskola*, introduced in 1963) and other two-year lines; vocational school; and direct labor-market entry. The categories of the age-16 parent item analyzed in Section 5.3 mirror this menu exactly (start working; practical traineeship; a two-year line; the three–four-year gymnasium). In 1971 (precisely as the cohort left compulsory school) these forms were merged into the integrated upper-secondary school (*gymnasieskola*), within which two-year and three–four-year lines persisted. Educational attainment is measured on the six-level Swedish educational nomenclature (SUN) in the registers, so the parental categories map naturally onto the realized scale.

The labor market, 1990–2022. The registers observe the cohort's earnings and employment from 1990, when the cohort was around 35, to 2022, when it reached 67. Three institutional layers shape what those person-level summaries contain, and in particular their sex composition. First, the family-policy architecture of the Swedish welfare state (Esping-Andersen, 1990) changed exactly as this cohort formed families and entered the labor market: joint income taxation of spouses was replaced by separate, individual taxation in 1971, sharply strengthening married women's incentives to work; in 1974 the maternal insurance was replaced by a parental insurance covering both parents (the first paid parental-leave system open to fathers) and from 1995 one month of the leave was reserved for each parent (the "father's month"), extended to a second reserved month in 2002; and publicly provided childcare expanded rapidly through the 1970s and 1980s. These reforms underpin the high, near-universal female labor-force attachment visible in the cohort's employment shares (women 0.81, men 0.80; Appendix Table A1), while mid-career earnings continue to show

a large sex gap. Second, the earnings window contains three macroeconomic episodes: the early-1990s Swedish banking and currency crisis (Englund, 1999), which struck when the cohort was in its late thirties (the first years of the register window); the 2008–2009 global financial crisis in the cohort's early fifties; and the 2020–2021 pandemic in its mid-sixties. Because earnings are averaged over ages 45–54, they are measured largely in the 2000s and 2010s, between the first two episodes. Third, the pension system was reformed in the 1990s toward notional defined contributions with a flexible claiming age, which shapes the late-career employment margin that the employment share and the age-67 observation window capture.

3.3 Variables

The parent-reported preference and expectation. Two parent-reported items measured at age 13 form the core pair of regressors. The first records the orientation the parents *preferred* for the child (vocational versus academic orientation); the academic-preference indicator equals one when the parents preferred the academic orientation. The second records the orientation the parents *expected* the child would choose; the academic-expectation indicator equals one when the parents expected the child would take the academic orientation. Preference is observed for 1,030 cohort members and expectation for 1,028; 71 percent of parents preferred the academic orientation and 63 percent expected the child would take it. A coherence check against a separately worded education-preference item (reverse-coded) confirms the polarity of the preference item: the two agree for 0.930 of the 1,001 members for whom both are observed (computed in the sample-construction script from the analytic extract).

Cross-classifying preference and expectation defines four groups: the *both-academic* group, who both preferred and expected the academic orientation (628); the *both-vocational* group, who preferred and expected the vocational orientation (282); the *preference-only* group, who preferred the academic orientation but expected the vocational one (92); and a small *expectation-only* group, who expected the academic orientation but did not prefer it (13). The two items are positively but far from perfectly correlated ($r = 0.78$; they agree for 90 percent of members). What distinguishes them empirically is a thin 105-member off-diagonal (the 92 preference-only and 13 expectation-only members), which is also why, though each measure is individually associated with its own margin, a formal test has limited power to separate the two.

Supporting parental and pupil items. The descriptive tables add contextual items: whether parents discussed orientation with the child at 13; whether they judged an academic

education to suit the child and to pay off; and three age-16 items (how much parents discussed the child's future, whether they judged the child suited to a long academic education, and whether parent and pupil agreed on the school type). The two ordinal age-16 items are re-oriented at construction so that higher values mean more of the named construct (more discussion; judged more suited); the raw codings run in the opposite direction and the reversal formulas are documented in the construction script and in Appendix B.

Two age-16 pupil items record whether the child wanted what the mother, or the father, wanted; these underlie the exploratory compliance analysis and are observed for only about 430 cohort members.

Realized attainment and labor-market outcomes. The registers observe members annually, so labor-market outcomes are constructed from person-year data and stated with time indices. Educational attainment is the highest education level attained between 1990 and 2022, coded on the six-level Swedish education scale (S_i, 1–6, higher = more education), with a post-secondary indicator (C_i, one when the level is at least four). Earnings enter in two forms. The main wage outcome is the person-level average of annual wage income over the prime-age window,

$$\bar{w}_i \;=\; \frac{1}{|T_i^w|}\sum_{t\in T_i^w} w_{it}, \qquad T_i^w = \{t:\ \text{age}_{it}\in[45,54],\ w_{it}\ \text{observed}\}, \tag{7}$$

in thousands of Swedish kronor.

The ages 45–54 window is a deliberate choice, not a convenience. Although the registers hold annual earnings for 1990–2022, a single-year figure is a biased proxy for lifetime earnings, and the bias moves systematically over the life cycle (Haider and Solon, 2006). Measured in the cohort's thirties, earnings are depressed precisely for members who stayed in school longest and entered the labor market latest (an education-timing bias that would understate the very schooling gradient this paper studies); measured in the sixties, earnings absorb early-retirement and part-time transitions. The association between current and lifetime earnings is closest to one in the mid-forties to early fifties, so a mid-career average is the least-biased available proxy for permanent earnings; the same life-cycle pattern holds in Swedish register data (Böhlmark and Lindquist, 2006), and averaging over ten years additionally smooths transitory annual shocks. That the headline is not an artifact of this window is shown in Section 6: preference and expectation coefficients are stable across alternative windows built from the same annual data (Appendix Table A18). The panel specifications of Section 6 use the annual wage w_{it} itself.

Labor-force attachment is the person-level average of an annual employment indicator,

the share of register years 1990–2022 in which the member was employed, $\bar{e}_i = |T_i|^{-1} \sum_{t \in T_i} e_{it}$ with $e_{it} \in \{0, 1\}$, and e_{it} itself in the panel specifications. The averaged and person-year forms provide robustness checks on one another.

Subjective well-being at 67. The retirement-age questionnaire provides life satisfaction on the five-item Satisfaction With Life Scale (SWLS, summed 5–35, higher = more satisfied) (Diener et al., 1985; Pavot and Diener, 1993); the single-item OECD life-satisfaction question (0–10, higher = more satisfied); a three-item loneliness scale (UCLA-3, summed 3–9, higher = lonelier); and depressive symptoms from the depression subscale of a standard anxiety-and-depression instrument (HAD-D, summed 0–21, higher = more depressive symptoms; items are pre-oriented at source and summed without reversal). All four scales are used in their documented orientations. Loneliness and depressive symptoms run *against* well-being, so a positive coefficient on those outcomes indicates worse late-life well-being. Outcomes are observed for roughly 590–610 cohort members who responded at 67.

Control variables. The control vector is fixed across specifications and comprises sex; birth-year group entered as a three-level dummy ($\leq$1954, 1955, $\geq$1956, with 1955 as the reference); two indicators for the family's socioeconomic group (white-collar and self-employed or farm, with worker-group families as the reference), taken from the family's classification in the 1970 census (the family measure with the widest coverage of the cohort, including incomers); family education from the grade-three parent questionnaire of 1965; an indicator for an intact family at grade three; and a same-wave measure of cognition (the child's total score on the DBA intelligence battery administered at the same 1968 wave as the parental items, the sum of six sub-test stanines, range 6–54, higher = stronger performance). Including cognition is central to the design, so parent-reported measures are not merely proxying for the child's contemporaneously measured ability.

Two scale conventions matter for reading the descriptive tables. The census socioeconomic classification and the family-education scale are both coded 1 = highest in the raw sources; both are re-oriented at construction ($x' = 8 - x$) so that higher values mean higher standing and more education, and census codes outside the documented 1–7 frame are set to missing. Because the top of the re-oriented socioeconomic scale mixes self-employed and farm families, the scale is not strictly ordered by status at its top; the models therefore enter the family's socioeconomic group as two categorical indicators, and the 1–7 scale appears only descriptively. The census measure is recorded two years after the parental items; it is treated as a slow-moving background characteristic, and Section 4.1 returns to its timing. Following project convention, the analytic sample is defined by the outcome and primary

exposures; the actual analytic N is reported in every table.

3.4 Descriptive statistics

Appendix Table A1 reports the analytic variables for the full cohort and separately for women and men, with tests of sex differences and the observed count for each variable. The two parent-reported measures do not differ significantly by sex: 73 percent of girls' parents and 69 percent of boys' parents stated a preference for the academic orientation ($p = 0.147$), and expectations are similarly balanced (65 versus 61 percent, $p = 0.133$).

Realized outcomes do differ by sex in the familiar directions. Women attain more schooling (mean education level 3.69 versus 3.46, $p < 0.001$; post-secondary share 0.43 versus 0.35, $p = 0.004$) but earn substantially less at mid-career (mean wage 209.90 versus 290.47 thousand kronor, $p < 0.001$), while employment shares are indistinguishable (0.81 versus 0.80, $p = 0.700$). The well-being outcomes at 67 are similar across sexes, with only the single-item OECD measure marginally higher for men (7.81 versus 7.50, $p = 0.081$). Two age-16 pupil items show the only large sex gaps in the parental block: girls were markedly more likely to report wanting what their mother (0.76 versus 0.56, $p < 0.001$) or father (0.75 versus 0.62, $p = 0.006$) wanted.

Table A2 repeats the battery across the three birth-year groups. The age-13 measures are mechanically related to birth-year group because the off-modal groups are small and, in the $\geq$1956 group, degenerate (every observed parent in that small group preferred and expected the academic orientation). This degeneracy is the reason birth year enters the models only as a three-level dummy and is treated as a nuisance control rather than a substantive moderator. The substantively informative birth-year gradient is in realized outcomes: members born later attain more education, earn more, and report fewer depressive symptoms at 67, a pattern consistent with the strong birth-year gradient in measured cognition (23.76, 30.81, 38.53 summed stanines across the three groups, $p < 0.001$).

Table A3 repeats the battery by cohort-entry route: original grade-three members, later in-migrants ("incomers"), and a small "movers" group. The table documents that the parent questionnaire conditions on parental response: family-structure items are largely missing for incomers because they joined the cohort after the grade-three family measurement, so the incomer column is blank for family education and intact-family status. The parent- reported measures themselves do not differ significantly by entry route. This composition is recorded so the reader can see which cohort members contribute to the parent-questionnaire-based models.

Table A4 collects the binary items into a single contrast table across all three groupings

for compactness.

3.5 From theoretical model to the empirical specification

The objects of the model of Section 2 map to the measured variables as follows. The skill stock θ_i is measured by the child's score on the DBA intelligence battery, taken at the same 1968 wave as the parental items. Expectation E_i and preference P_i are the two 1968 parent-reported items; the parent's concrete post-compulsory expectation, collected at age 16, provides a category-level counterpart. Completed schooling S_i is the highest education level recorded in the registers (SUN 1–6), with post-secondary attainment C_i as its upper margin. Earnings are measured as annual wage income w_{it} from the registers, entering the analysis both averaged over ages 45–54 and as a person-year panel. The attachment margin is an indicator for being employed in year t.

The control vector $\mathbf{X}_i$ comprises the family's socioeconomic group, family education, intact family structure, sex, and birth-year group (Section 3.3). The valuation ν_i, investments I_i, and uncertified skills N_i are not observed. The valuation term is identified residually as preference variation conditional on expectation and cognition, and the presence of the direct channel λN_i is inferred from the conditional pattern H2 describes rather than measured directly. The regressions that implement these mappings are set out next.

4 Identification and econometric model

4.1 Identification strategy

Estimand and analyzed population. The source of variation is cross-sectional: differences across cohort members in parent-reported preference and expectation recorded at age 13. The estimand is the conditional association of each parent-reported measure with each adult outcome (the coefficients β_P and β_E of equation (9)) in the population for which the measures are *defined*, namely cohort members who were present in the cohort at the 1968 lower-secondary wave and whose parents returned the questionnaire.

This is a descriptive estimand, not a causal effect. Conditional-mean independence is *not* assumed: parent-reported measures are not randomly assigned, and unobserved determinants of attainment may be correlated with what parents hope for and expect. What the design *does* secure is that the preference coefficient is measured conditional on the expectation and same-wave cognition, so it is not mechanically driven by parents' accurate reading of the child's current ability.

Observation equation. The analytic samples arise from the 1,392 cumulative cohort members through a sequence of selections,

$$O_i \;=\; P_i \times Q_i \times R_i^{\text{reg}} \qquad \Big(\times\, S_i \times R_i^{67} \text{ for the well-being models}\Big), \tag{8}$$

where P_i indicates presence in the cohort at the 1968 (grade-six) wave, Q_i parental questionnaire response given presence, R_i^{reg} appearance in the post-1990 registers, S_i survival to the retirement-age wave, and R_i^{67} response at 67. The first margin is structural, not stochastic: the 22 "movers" who had left the town after grade three could not receive the 1968 parent questionnaire, so for them the item is undefined rather than missing. Among members present when the item was collected, it is observed for 869 of the 1,009 present from grade three, 122 of the 279 incomers who joined the cohort's classes later, and 39 of a residual group of 82 with a mixed entry record (1,030 members in all, matching the reported exposure count). Because the item is undefined for those outside the 1968 wave, the estimand is stated on the population where the measure exists rather than "corrected" toward the full cohort of 1,392 (Table A3 reports the three principal entry groups).

Threats and the analysis that addresses each. Each selection margin and each leading threat maps to a named analysis.

(i) *The measures proxy for the child's ability* (θ_i in the model): addressed by conditioning on same-wave cognition in every specification and, as a deliberately conservative over-adjustment, on post-exposure school grades at 16 (Section 6).

(ii) *Parental non-response within the 1968 wave* (the stochastic part of Q_i): characterized and corrected rather than assumed away. A balance table (Appendix Table A16) shows that responders and non-responders differ on *inputs*: responders are far less often incomers, are of somewhat lower census family standing, and score higher on the age-13 cognitive test, but do not differ significantly on the adult register *outcomes* (education and wage). Re-estimating the headline models with stabilized inverse-probability-of-response weights, from a logit response model on the cohort-wide register and census characteristics (Appendix Table A17), leaves the coefficients essentially unchanged: the expectation on education moves from 0.290 to 0.284 and the preference on wages from 39.8 to 40.1. Response selection on observables therefore does not drive the estimates; the weighting corrects observed selection only, so selection on unobservables (in particular a parental disposition to answer that is itself correlated with the preference–expectation gap, such as the motivation of a parent who prefers but does not expect the academic route) cannot be ruled out.

(iii) *Survival and response at 67* (S_i, R_i^{67}, affecting only the well-being models): addressed

by inverse-probability-of-response weights (Section 6) and by carrying the primary contrast on the larger register-outcome sample, which does not condition on either margin (Section 6).

(iv) *The census family measure post-dates the parental items by two years*: no correction is available; it is treated as a slow-moving background characteristic, and to the extent the 1970 classification already reflects the family's trajectory, that is absorbed by the family coefficients rather than by the preference and expectation coefficients, which additionally condition on cognition and on each other.

(v) *The four well-being outcomes are correlated*: the exploratory affective analysis carries the multiple-testing caution of Section 4.3. Analyses that answer no identification question (the descriptive batteries and expectation-by-preference profile) are labeled descriptive or exploratory where they appear.

4.2 Econometric model

For each register outcome Y_i (education level, post-secondary attainment, mean wage over ages 45–54, employment share), the analysis estimates a sequence of person-level regressions. The baseline is the additive controlled model,

$$Y_i = \alpha + \beta_P P_i + \beta_E E_i + \mathbf{X}_i'\boldsymbol{\gamma} + e_i, \tag{9}$$

where P_i and E_i are the age-13 preference and expectation indicators (each one when the parent preferred, or expected, the academic orientation). The control vector $\mathbf{X}_i$ is fixed across specifications (Section 3.3): sex, birth-year group, the two family-socioeconomic-group indicators, family education, intact family structure, and same-wave cognition. The coefficient β_E measures the association of expectation with the outcome holding preference and controls fixed; β_P measures the association of preference holding expectation and controls fixed. Because cognition is among the controls, β_P is preference net of the child's measured ability. H1 and H2 are read off β_E and β_P in the schooling and earnings equations.

Two variants of equation (9) sharpen the tests. To test the complementarity of H3, the product of the two indicators is added,

$$Y_i = \alpha + \beta_P P_i + \beta_E E_i + \beta_{EP} E_i P_i + \mathbf{X}_i'\boldsymbol{\gamma} + e_i, \tag{10}$$

where $\beta_{EP} > 0$ if the association of one measure is larger when the other also points to the academic orientation. To test the outcome-specific split of H1 and H2 in a single coefficient,

the two indicators are recombined into their sum and their difference,

$$Y_i = \alpha + \beta_O\,(E_i + P_i) + \beta_G\,(E_i - P_i) + \mathbf{X}_i'\boldsymbol{\gamma} + e_i. \tag{11}$$

Here $E_i + P_i$ is the family's *academic orientation*: it equals 0 when neither preference nor expectation points to the academic orientation, 1 when exactly one does, and 2 when both do; its coefficient is β_O. The difference $E_i - P_i$ is the *expectation-minus-preference gap*: it equals -1 when the family prefers the academic orientation but does not expect it, 0 when preference and expectation agree (both point to the academic orientation or both to the vocational one), and $+1$ when the family expects the academic orientation but does not prefer it; its coefficient is β_G. Equation (11) fits identically to equation (9), and by construction $\beta_G = (\beta_E - \beta_P)/2$. The sign and significance of β_G therefore say directly whether expectation and preference diverge on a given outcome, and the outcome-specific split implies $\beta_G > 0$ for schooling and $\beta_G < 0$ for earnings.

Two further specifications complete the design. Before any controls are added, a naive ladder enters the two measures without covariates, one at a time, then together, then with their interaction:

$$Y_i = \alpha_1 + \beta_P\,P_i + e_{1i}, \tag{12}$$

$$Y_i = \alpha_2 + \beta_E\,E_i + e_{2i}, \tag{13}$$

$$Y_i = \alpha_3 + \beta_P\,P_i + \beta_E\,E_i + e_{3i}, \tag{14}$$

$$Y_i = \alpha_4 + \beta_P\,P_i + \beta_E\,E_i + \beta_{EP}\,E_iP_i + e_{4i}, \tag{15}$$

Each equation is estimated on the widest sample its own variables allow (complete case per specification, $N \approx 960$–994), so the ladder separates raw associations from the role of controls. Equations (9)–(11) are estimated on the complete-case sample with every control observed.

For the earnings outcome, one further specification adds completed schooling to equation (9) (the transmission check of H2, which asks how much of each measure's earnings association passes through the diploma (Section 6)).

Finally, because the register outcomes are built from person-year data, wage and employment results are re-estimated on the person-year panel itself,

$$y_{it} = \lambda_t + \beta_P\,P_i + \beta_E\,E_i + \mathbf{X}_i'\boldsymbol{\gamma} + u_{it}, \tag{16}$$

where y_{it} is the annual wage at ages 45–54 or the annual employment indicator over 1990–2022, λ_t are year fixed effects, P_i, E_i, and $\mathbf{X}_i$ are as in equation (9), and u_{it} is the within-

person error. Standard errors are clustered on the person (Section 6).

A person fixed-effects estimator is not available here: P_i, E_i, and $\mathbf{X}_i$ are time-invariant, and the within transformation removes them. Equation (16) is what the panel identifies for time-invariant regressors: pooled estimation that uses every annual observation, absorbs common year shocks, and accounts for within-person dependence in the errors.

Data structure and estimators. The unit of analysis is the person, not the person-year. The register outcomes are person-level summaries of the 1990–2022 panel (the highest education level attained, mean wage income over ages 45–54, the share of register years employed), and the two exposures are measured once, at age 13, and never vary within person. Summarizing labor-market outcomes as means over the mid-career window (wages over ages 45–54, employment over 1990–2022) rather than year by year is deliberate: the average damps transitory, year-specific earnings shocks and isolates the persistent mid-career earnings capacity that the Mincer relation concerns.

Panel estimators (person-year fixed effects and related within transformations) are not applicable to the exposures: a within-person transformation cannot identify a coefficient on a time-invariant regressor, and the summary outcomes carry no within-person dimension. Every model in the paper is therefore a person-level cross-sectional regression; the person-year panel with year fixed effects and person-clustered errors (equation (16), Appendix Table A12) is reported only as a robustness check on averaging, and it reproduces the averaged estimates.

Estimators are, by outcome type: (i) *continuous outcomes* (education level treated cardinally (SUN 1–6), mean wage (1,000 SEK), employment share (0–1), and the four well-being scales), estimated by *ordinary least squares*; (ii) the *binary* post-secondary indicator estimated as a *linear probability model* by OLS, the project's primary convention because coefficients read directly in probability points; (iii) all least-squares specifications carry *HC3 heteroskedasticity-robust standard errors* (MacKinnon and White, 1985); (iv) as confirmation that attainment results are not artifacts of the linear functional form, Section 6 re-estimates the education-level model as an *ordered logit* and the post-secondary model as a *probit* with average marginal effects; (v) response-weighted sensitivity re-estimates the well-being models by *weighted least squares* with stabilized inverse-probability-of-response weights. This is estimation of conditional associations, not out-of-sample prediction.

The same equation is re-used with different right-hand-side codings. The group specification replaces the two indicators with the four expectation-by-preference group indicators (reference: both-vocational), so each group is compared to families aligned on the vocational orientation. The well-being specification replaces the two indicators with an indicator for

an *unmet* academic preference (parents preferred the academic orientation but the child did not reach post-secondary attainment) and a *realized* academic preference (preferred and reached it), with the reference category being no academic preference. Step-by-step covariate additions are not written as separate equations; the control vector is held fixed across headline specifications, and alternative covariate sets appear as separate robustness columns (Section 6).

4.3 Confirmatory versus exploratory analyses

One contrast is confirmatory and pre-specified: the preference-versus-expectation contrast of H1 and H2, with each parent-reported measure holding the other and the controls fixed. Each hypothesis of Section 2.3 maps to named estimates. H1 is tested on the schooling columns of equation (9), with ordered-logit and probit models as functional-form confirmation. H2 is tested on the earnings column of equation (9) together with the schooling-added specification and the person-year panel of equation (16). The outcome-specific split of H1 and H2 is also tested directly by the expectation-minus-preference gap coefficient of equation (11). H3 is tested by equation (10) and the interaction column of the naive ladder.

The confirmatory contrast is reported and interpreted directly. All other results (the four-group profile and contrasts, the category-level expectation map, the late-life well-being models, and the age-16 compliance interaction) are exploratory or descriptive. They are reported honestly and in full, but their nominal p-values are read against the usual multiplicity caution: a wide battery-association screen of cohort measures returns hundreds of significant correlations after a Benjamini–Hochberg false-discovery-rate adjustment (Benjamini and Hochberg, 1995) (Appendix Figure A1; 1,708 tests, 540 significant at $q < 0.05$), so a single nominally significant exploratory coefficient is weak evidence on its own.

5 Results

5.1 What each parent-reported measure carries on its own

Tables 1 and 2 report the naive ladder of equations (12)–(15) (each measure alone, both together, and both with their interaction, without any controls) for the two schooling outcomes and the two labor-market outcomes.

Table 1. Naive specification ladder: schooling outcomes

	Education level (SUN 1–6)				Post-secondary (LPM)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Parent's preference: academic orientation	0.937*** (0.065)		0.519*** (0.106)	0.448*** (0.119)	0.399*** (0.027)		0.212*** (0.047)	0.160*** (0.053)
Parent's expectation: academic orientation		0.896*** (0.065)	0.518*** (0.104)	0.139 (0.192)		0.385*** (0.028)	0.231*** (0.048)	-0.047 (0.083)
Preference × expectation				0.440** (0.224)				0.323*** (0.098)
N	992	990	977	977	992	990	977	977

Note: Ordinary least squares (columns under "Post-secondary" are linear probability models) with HC3 robust standard errors in parentheses; no controls. Columns: (1) preference only, equation (12); (2) expectation only, equation (13); (3) both, equation (14); (4) both plus interaction, equation (15). Each column uses the widest complete-case sample its variables allow. Education is the highest register level 1990–2022 (SUN 1–6, higher = more education). Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table 2. Naive specification ladder: labor-market outcomes

	Mean wage 45–54 (1,000 SEK)				Employment share 1990–2022			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Parent's preference: academic orientation	105.468*** (11.393)		53.182*** (15.624)	40.126** (16.822)	0.093*** (0.018)		0.046 (0.032)	0.029 (0.034)
Parent's expectation: academic orientation		101.968*** (11.904)	64.317*** (16.759)	-4.793 (38.700)		0.089*** (0.017)	0.057* (0.029)	-0.037 (0.090)
Preference × expectation				80.403* (42.813)				0.109 (0.095)
N	975	973	960	960	994	992	979	979

Note: As Table 1. Wage is mean annual wage income over ages 45–54 (1,000 SEK); the employment share is the share of register years 1990–2022 employed (0–1). Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Three facts stand out. First, *one measure at a time*, preference and expectation are nearly interchangeable. Alone, a parental preference for the academic orientation comes with 0.937 more education levels and an expectation with 0.896 (both $p < 0.01$, $N = 992$ and 990); on mid-career wages the two raw associations are 105.5 and 102.0 thousand kronor. This is what the framework predicts for two items that share the same information core.

Second, *entered together* (column 3), the two measures split the schooling association almost exactly evenly (0.519 and 0.518, both $p < 0.01$, $N = 977$), and both remain associated with wages (53.2 and 64.3, both $p < 0.01$). On the employment share, expectation is marginal (0.057, $p = 0.051$) and preference is not statistically significant (0.046, $p = 0.147$). The outcome-specific split is therefore *not* visible in the raw data; it emerges conditional on family background and the child's cognition (Section 5.2).

Third, the *interaction* column shows that raw gradients are carried mainly by families aligned on the academic orientation. The expectation-by-preference interaction is positive for the education level (0.440, $p = 0.050$), strongly so for post-secondary attainment (0.323, $p = 0.001$), and suggestive for wages (80.4, $p = 0.060$). The expectation main effect in this saturated column compares the 13-member expectation-only cell with the both-vocational group and is correspondingly imprecise. The interaction specification is the regression form of the four-group profile reported in Appendix Tables A6 and A7.

The ladder also answers a sample question that the controlled model alone cannot. Moving from column (3) ($N = 977$) to the fully controlled specification ($N = 705$) loses roughly 270 members to control missingness. Re-estimating column (3) on exactly the controlled sample leaves the qualitative ordering intact and already visible *before* any adjustment: expectation leads on the education level (0.609 versus 0.432 for preference, both $p < 0.01$), while preference leads on wages (66.2 versus 51.2, both $p < 0.01$) and carries the employment share (0.071, $p = 0.026$, versus 0.019, $p = 0.498$). The movement from the naive to the controlled estimates therefore reflects covariate adjustment (family background and cognition absorbing the shared information core), rather than the change of sample.

5.2 The pre-specified contrast: expectation on schooling, preference on earnings

The headline result is in Table 3. Panel A estimates the additive controlled model, equation (9), for the four register outcomes with the two parent-reported measures entered together; the pre-specified contrast of H1 and H2 is read across its first two rows. Panel B adds the expectation-by-preference interaction, equation (10), and tests H3. Panel C reports the orientation-and-gap reparametrization, equation (11), whose gap coefficient tests directly

whether the two measures can be told apart.

For **educational attainment**, parental *expectation* that the child would take the academic orientation is associated with 0.290 of a level on the six-level education scale (HC3 standard error 0.118, $p = 0.014$) and with a 0.126 higher probability of post-secondary attainment (0.056, $p = 0.025$). Parental *preference*, holding expectation fixed, is only weakly associated with schooling: 0.203 of a level (0.122, $p = 0.097$) and 0.088 on the post-secondary margin (0.057, $p = 0.127$), neither reaching the 5 percent level. This is the pattern H1 states: expectation, not preference, tracks the diploma.

For **earnings**, the pattern reverses. Parental *preference* is associated with 39.782 thousand kronor more in mean annual wage income over ages 45–54 (19.179, $p = 0.038$), while parental *expectation* is unrelated to earnings (17.291, 18.747, $p = 0.356$). Preference is also associated with a 0.053 higher employment share (0.032, $p = 0.094$), again where expectation is essentially zero (-0.001, 0.029, $p = 0.973$). This is the pattern H2 states: holding expectation fixed, preference is associated with a labor-market margin rather than with schooling.

Taken together, the two rows of Panel A trace the outcome-specific split: each hypothesis holds on its own margin, i.e., expectation is significant on schooling (H1) and preference on earnings (H2). Whether the two measures differ from each other is a separate question, tested directly by the gap coefficient. Figure 1 displays the same two coefficient rows with 95 percent confidence intervals, one panel per outcome in its natural units. The expectation interval clears zero on the two post-secondary margins and straddles it on both labor-market margins, while the preference interval clears zero on earnings and straddles it on both post-secondary margins. The point estimates are economically meaningful: roughly a third of an education level and 13 percentage points of post-secondary attainment for expectation, and roughly 40,000 kronor of annual mid-career earnings for preference. The control coefficients (Appendix Table A5) behave as expected (family education and same-wave cognition are strongly associated with attainment, and women earn markedly less at mid-career conditional on attainment), which lends face validity to the specification.

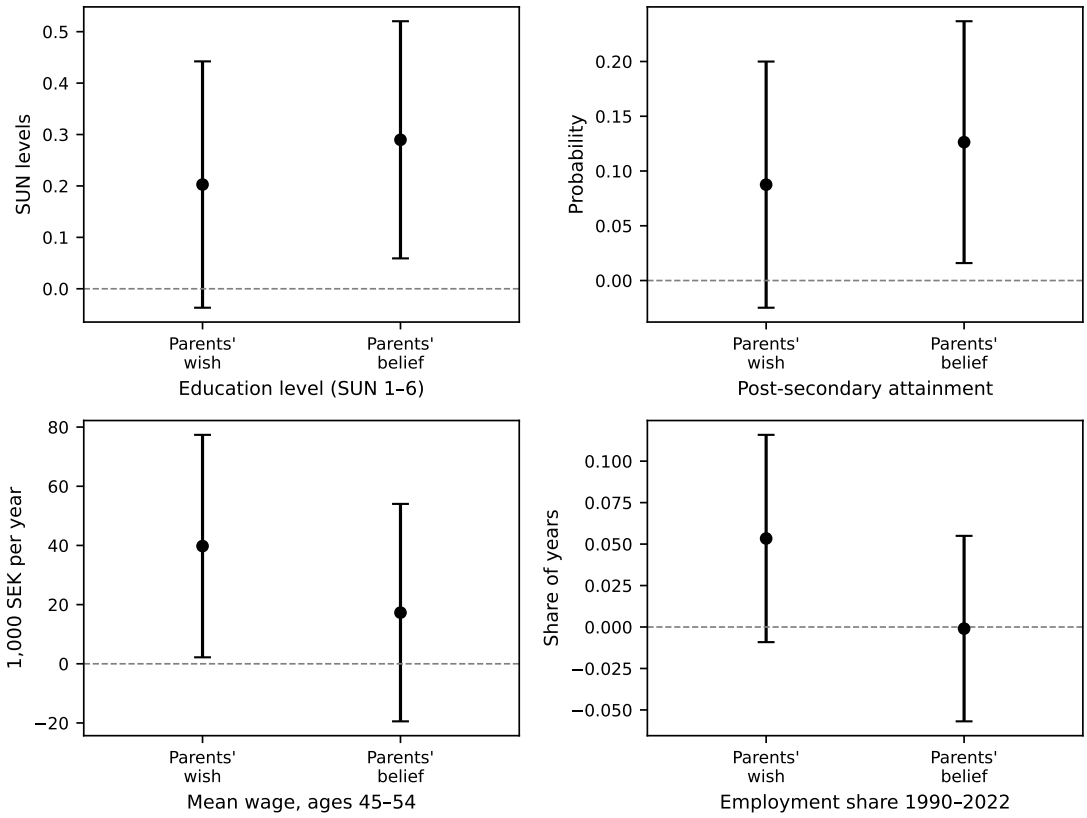

Figure 1. The preference and the expectation across the four register outcomes

Note: Point estimates and 95 percent confidence intervals for the parent's preference and the parent's expectation from the headline specification of Table 3 (equation 9, HC3 standard errors), one panel per register outcome in its natural units. Intervals are the coefficient ± 1.96 standard errors as reported in Table 3. The expectation interval clears zero only on the two post-secondary margins; the preference interval clears zero only on earnings.

Table 3. Controlled estimates: the parent-reported preference and expectation at age 13 on adult register outcomes

	Education (SUN)	Post-sec. (LPM)	Wage 45–54	Emp. share
Panel A: additive specification				
Parent's preference: academic orientation	0.203* (0.122)	0.088 (0.057)	39.782** (19.179)	0.053* (0.032)
Parent's expectation: academic orientation	0.290** (0.118)	0.126** (0.056)	17.291 (18.747)	-0.001 (0.029)
N	705	705	699	707
Panel B: with the preference-by-expectation interaction				
Parent's preference: academic orientation	0.129 (0.139)	0.028 (0.062)	26.078 (20.655)	0.049 (0.035)
Parent's expectation: academic orientation	-0.054 (0.220)	-0.151 (0.108)	-45.553 (47.418)	-0.022 (0.087)
Preference × expectation	0.406 (0.253)	0.328*** (0.123)	74.525 (51.475)	0.025 (0.090)
N	705	705	699	707
Panel C: academic orientation and expectation-minus-preference gap				
Academic orientation ($E_i + P_i$: 0/1/2)	0.246*** (0.051)	0.107*** (0.022)	28.536*** (7.474)	0.026** (0.012)
Expectation−preference gap ($E_i - P_i$: −1/0/+1)	0.044 (0.109)	0.019 (0.052)	-11.246 (17.429)	-0.027 (0.028)
N	705	705	699	707

Note: Ordinary least squares estimates of equations (9) (Panel A), (10) (Panel B), and (11) (Panel C) with HC3 robust standard errors in parentheses. The preference (expectation) indicator equals one when the parent preferred (expected) the child would take the academic (gymnasium-preparatory) orientation rather than the vocational one (Section 3.2). In Panel C, "Academic orientation" is the sum $E_i + P_i$ (0 when neither the preference nor the expectation points to the academic orientation, 1 when exactly one does, 2 when both do) and "Expectation−preference gap" is the difference $E_i - P_i$ (-1 when the family prefers the academic orientation but does not expect it, 0 when the preference and expectation agree, $+1$ when it expects but does not prefer it); Panel C fits identically to Panel A, and the gap coefficient equals $(\beta_E - \beta_P)/2$. Education (SUN) is the highest education level 1990–2022 on the six-level Swedish scale (higher = more education); Post-sec. is a linear probability model for post-secondary attainment (level ≥ 4); Wage 45–54 is mean annual wage income over ages 45–54 (1,000 SEK); Emp. share is the employment share 1990–2022. All columns condition on sex, birth-year group ($\leq$1954 / 1955 / $\geq$1956), the two family-socioeconomic-group indicators (worker-group reference), family education, an intact family at grade three, and same-wave cognition at age 13. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***). The contrast across the first two rows of Panel A (expectation on schooling, preference on earnings) is the paper's pre-specified primary result (H1 and H2), displayed in Figure 1; Panel B tests H3; Panel C's gap coefficient tests the outcome-specific split directly. Full coefficient sets for Panel A are in Appendix Table A5.

Panel B tests H3. The expectation-by-preference interaction is positive on every outcome, as the complementarity hypothesis states, and statistically significant on the post-secondary margin (0.328, $p = 0.008$). On the education level (0.406, $p = 0.109$) and on wages (74.5, $p = 0.148$) it is positive but not significant at conventional levels, and on the employment share it is essentially zero. In the interacted specification, the main effects compare the small non-aligned cells (the 92-member preference-only and 13-member expectation-only groups) with the both-vocational reference and are correspondingly imprecise. H3 is thus supported where the data are best able to test it (the diploma margin) and not rejected elsewhere.

Panel C reads the same contrast the way H1 and H2 state it most directly. The *academic orientation* $E_i + P_i$ (how far toward the academic orientation family reports point overall) is strongly and significantly associated with every outcome (0.246 education levels, 0.107 post-secondary, 28.5 thousand kronor, 0.026 employment share, all $p < 0.05$). The *gap* $E_i - P_i$ is the direct test of whether the two measures differ, with coefficient equal to $(\beta_E - \beta_P)/2$. Its sign matches the framework on each margin (positive where expectation leads (education +0.044, post-secondary +0.019) and negative where preference leads (wages −11.2, employment −0.027)), so the single gap coefficient changes sign between schooling and earnings as H1 and H2 jointly imply. It is not, however, statistically distinguishable from zero on any one outcome (p between 0.33 and 0.71).

The outcome-specific split therefore rests on the pattern (expectation significant on schooling and silent on earnings, preference the mirror image, the gap coefficient flipping sign across margins, and the schooling-conditioned transmission of Section 6), not on a knife-edge difference between two coefficients on any single margin.

A joint test across the two margins does not overturn this reading. Stacking the standardized schooling and earnings equations and asking whether the gap differs between them (the single most demanding form of the claim) gives a cross-margin difference of −0.078 standard deviations (person-clustered $p = 0.44$), and a joint Wald test that the gap is zero on both margins does not reject ($\chi^2_2 = 0.70$, $p = 0.71$; Appendix Table A15). Point estimates line up the same way on every specification, but no test (per-margin or joint) rejects that the two measures relate to schooling and earnings in the same way. The outcome-specific split is therefore reported throughout as a directional regularity, not as an established difference between the two measures.

5.3 Parental expectations and adult reality

The measures analyzed so far are binary. At age 16, the parents were also asked a concrete, category-level version of the same question: what the pupil would do on finishing compulsory

school, with answers spanning the post-compulsory menu of Section 3.2: start working, a practical traineeship, a two-year gymnasium line, another two-year line, the three–four-year gymnasium, or something else ("don't know" responses are set to missing). Table 4 maps each expectation category onto the education levels children actually reached in the registers, for the 881 cohort members with both an expectation stated and a register education record. This is the category-level comparison of parental expectation and adult reality. It is descriptive, and it can be read as a disaggregated view of the expectation content that H1 summarizes.

Table 4. Parent's post-compulsory expectation at age 16 and the child's realized register education

		Realized education level (SUN), row share					
Parent's expectation at 16	N	≤2	3	4	≥5	Mean SUN	Post-sec.
Start working	44	0.50	0.43	0.00	0.07	2.66	0.07
Practical traineeship	72	0.35	0.58	0.01	0.06	2.78	0.07
Two-year gymnasium line	284	0.07	0.57	0.06	0.30	3.60	0.36
Other two-year line	131	0.16	0.61	0.02	0.21	3.28	0.23
Three–four-year gymnasium	336	0.02	0.23	0.15	0.60	4.36	0.75
Something else	14	0.14	0.64	0.07	0.14	3.14	0.21
All with an expectation stated	881	0.11	0.44	0.08	0.37	3.72	0.45

Note: Rows are the categories of the age-16 parent item "What shall the pupil do on finishing compulsory school?" (1971 parent questionnaire; "don't know" set to missing). Columns give the distribution of the child's realized highest education level 1990–2022 on the six-level SUN scale (row shares), the row mean of the SUN level, and the row share reaching post-secondary attainment (SUN $\geq$ 4). Descriptive; $N = 881$ cohort members with both an expectation stated and a register education record.

The gradient is steep and monotone in the academic ambition of the expectation. Where parents expected the child to start working, the children's mean realized level is 2.66 and 7 percent reached post-secondary education; where parents expected a practical traineeship, 2.78 and 7 percent; where they expected a two-year gymnasium line, 3.60 and 36 percent; another two-year line, 3.28 and 23 percent; and where they expected the three–four-year gymnasium, 4.36 and 75 percent. The small residual category ("something else," 14 members) sits near the two-year groups.

Read against the framework, the category map is the disaggregated content of the expectation–schooling association: parents' concrete expectations order children's realized education levels almost perfectly, half a century before the register window closes.

5.4 The four expectation-by-preference groups

The two-indicator specification imposes additivity on preference and expectation. Appendix Table A6 relaxes that by profiling the four expectation-by-preference groups descriptively, and Appendix Table A7 estimates the same outcomes with the four groups entered as indicators (reference: both-vocational families). These analyses are exploratory; their regression form is the interaction column of the naive ladder (Section 5.1).

The descriptive profile (Appendix Table A6) shows a graded structure. Both-academic families (628 members) are the most advantaged on every margin: family education 3.34, same-wave cognition 33.88, and 0.56 post-secondary attainment, with mean mid-career earnings of 296.72 thousand kronor. Both-vocational families (282) are the least advantaged: family education 2.00, cognition 24.13, 0.12 post-secondary, and earnings 180.98. The preference-only group (92) (parents who preferred the academic orientation but expected the vocational one) sits between the two aligned groups on essentially every dimension: family education 2.51, cognition 27.87, 0.28 post-secondary, and earnings 221.11. The small expectation-only group (13) is too sparse to interpret. The preference-only group is thus not a random perturbation of the both-vocational group; it is intermediate, consistent with preference carrying information that the both-vocational families lack.

The group regression (Appendix Table A7) confirms that the both-academic group drives the attainment gradient. Relative to both-vocational families, both-academic families attain 0.482 more education levels (0.101, $p < 0.01$), 0.205 more post-secondary probability (0.044, $p < 0.01$), and 55.050 thousand kronor more in earnings (14.898, $p < 0.01$). The preference-only group's coefficients are positive on every outcome but individually insignificant (0.129 education levels, $p = 0.352$; 26.078 thousand kronor of earnings, $p = 0.207$), which is unsurprising given the group's size; their sign and ordering nonetheless match the descriptive profile. These group estimates are consistent with (but, given their exploratory status and the small preference-only cell, do not independently establish) the headline outcome-specific split.

5.5 A secondary check: no subjective "scar" from an unmet academic preference

The exploratory affective implication of Section 2 is checked on the roughly 340–348 cohort members who responded at 67 by estimating equation (9) for the four well-being outcomes with unmet- and realized-academic-preference indicators replacing preference and expectation (reference: no academic preference). Estimates are reported in Appendix Table A8.

Neither an unmet nor a realized academic preference is associated with life satisfaction

(SWLS 0.898, $p = 0.357$, and 0.946, $p = 0.315$), the single-item OECD measure (both 0.023, $p > 0.94$), or loneliness (0.145, $p = 0.524$, and -0.039, $p = 0.862$). The only nominally significant coefficient (realized preference on the HAD depression scale, -0.734 ($p = 0.065$)) is at most marginal and, if anything, points toward slightly better mental health when the preference was realized.

The reading is a "good-news null": an unmet academic preference leaves no measurable affective scar at retirement age, contrary to the unmet-preference implication. This analysis is secondary to the paper's attainment question; its selection and power caveats are presented in Sections 6 and 7.

6 Robustness and sensitivity analyses

The robustness analyses map onto the three identification threats of Section 4.1: that the measures proxy for ability, that the well-being sample is selected, and that the exploratory affective tests are subject to multiplicity.

Adding school grades at age 16. Table A9 re-estimates the two attainment columns of the headline table after adding the child's mean school grades at age 16 (a post-exposure measure of realized school performance). This is a demanding check: grades lie causally between the age-13 items and adult attainment, so they will absorb measure content that operates through later school performance. Under this over-control, the expectation coefficient on education falls from 0.290 to 0.187 and is no longer significant (0.120, $p = 0.120$), and the preference coefficient is similarly attenuated (0.175, $p = 0.156$); grades themselves are very strongly associated with attainment (0.569 per grade point, $p < 0.01$).

The attenuation is expected and does not overturn the headline reading: it suggests that a large part of the expectation–schooling association runs through the school performance the child went on to achieve, which is itself part of the expectation's content rather than a confounder. This is a conservative over-adjustment and is reported as such; the headline specification (Table 3) deliberately does *not* condition on this post-exposure variable.

Ordered logit and probit estimates. Education level is an ordered six-level scale and the post-secondary indicator is binary, so the linear estimators of Table 3 are confirmed under corresponding response models on the same sample and right-hand side (Appendix Table A10). The *ordered logit* for education reproduces the asymmetry of the linear model: the expectation coefficient is 0.598 in log-odds ($p = 0.020$), while preference does not clear

the 5 percent level (0.477, $p = 0.081$). The implied average marginal effects on reaching at least post-secondary level are 0.118 for expectation and 0.092 for preference.

The *probit* for the post-secondary indicator puts expectation at $p = 0.048$ with an average marginal effect of 0.102 and preference at a statistically marginal $p = 0.055$ with an average marginal effect of 0.113, somewhat sharper for preference than the linear probability model's $p = 0.127$. The expectation–schooling half of the outcome-specific split is therefore robust to functional form. The statement that preference is silent on schooling is sharpest on the six-level scale and reads as marginal, rather than clearly null, on the binary margin. The earnings half of the contrast is unaffected (earnings and the employment share are continuous and remain least-squares).

Adding completed schooling to the earnings equation. Appendix Table A11 adds completed schooling to the controlled earnings equation. Schooling carries 48.3 thousand kronor per education level ($p < 0.01$), the Mincer margin.

Conditional on schooling, the expectation's earnings association collapses from 17.3 to 4.3 thousand kronor ($p = 0.808$): the diploma transmits essentially all of the expectation's earnings content, exactly as equation (3) implies for a measure whose only earnings channel is schooling.

The preference's association falls from 39.8 to 28.8 thousand kronor and is no longer statistically significant at conventional levels ($p = 0.110$). A positive beyond-schooling component remains in the point estimate, but the sample cannot establish it at the 5 percent level, and its magnitude (and hence any split of the total into schooling and beyond-schooling parts) is estimated too imprecisely to quantify.

H2's transmission statement is supported in sign and ordering: the expectation's earnings content travels through the diploma, and the preference retains a positive but imprecise component beyond it.

Person-year panel estimates. The averaged labor-market outcomes could in principle mask patterns in the underlying annual data. Appendix Table A12 therefore estimates equation (16) on the person-year panel itself, with year fixed effects and standard errors clustered on the person.

The panel estimates reproduce the person-level ones almost exactly. On annual wages at ages 45–54 (6,901 person-years, 699 persons), the preference coefficient is 38.6 thousand kronor ($p = 0.043$) and the expectation's 17.4 ($p = 0.347$), against 39.8 and 17.3 in the averaged model. On the annual employment indicator over 1990–2022 (49,187 person-years, 728 persons), preference carries 2.6 percentage points ($p = 0.077$) and expectation zero

($p = 0.985$). Averaging over years neither creates nor hides the outcome-specific split.

Alternative earnings windows. The mid-career window is chosen on life-cycle-bias grounds (Section 3.3), and the earnings result is consistent with, not an artifact of, that choice. Appendix Table A18 recomputes the additive wage model with earnings averaged over ages 40–49, 45–54, 50–59, 55–64, 40–59, and over all observed register years. The preference coefficient is positive and at least marginally significant in every window (from 26.2 to 43.0 thousand kronor, p between 0.038 and 0.091), while the expectation coefficient is significant in none (11–17 thousand kronor, $p \geq 0.23$).

The preference association traces a hump over the life cycle: weakest when the earliest years are mixed in (ages 40–49, or all years, where the more-educated were still entering work), strongest across the mid-career decade (ages 45–59), and easing again as the cohort nears retirement (ages 55–64). This is exactly the shape the life-cycle-bias argument implies, with the ages 45–54 headline near the top of the profile. The contrast between the two measures is therefore a property of the earnings data, not of any one window.

Inverse probability weighting of the well-being models. Table A13 re-estimates the four late-life well-being models with inverse-probability-of-response weights estimated for the retirement-age wave (Horvitz and Thompson, 1952; Wooldridge, 2007). The weighted estimates are close to the unweighted estimates of Table A8. The unmet- and realized-preference coefficients on the SWLS remain insignificant (0.903, $p = 0.430$, and 1.152, $p = 0.299$), and the marginal protective association of a realized preference with depressive symptoms persists (-0.928, $p = 0.062$). The good-news null is thus not an artifact of selective response at 67.

Construction of the response weights. The weights of Table A13 are register-based stabilized weights for inverse probability of response. They are constructed in a companion non-response analysis of this cohort by modeling the probability of responding at the retirement-age wave on observed register and childhood characteristics available for the entire cohort, and stabilizing each weight by the marginal response rate to limit the influence of extreme values (Horvitz and Thompson, 1952; Wooldridge, 2007). They enter the four well-being models through weighted least squares, up-weighting respondents who resemble non-respondents on those observables.

The complete-case specification (Table A8) remains the *primary* estimate. The weighted specification is a documented sensitivity that supplements rather than replaces it: response weighting corrects only the part of selection explained by observed characteristics, and the

response model is itself estimated, so a weighted coefficient is not granted headline status.

The register sample and the survey sample. The paper rests on two analytic samples that are selected differently, and both are reported rather than collapsed. The headline preference-versus-expectation contrast on attainment and earnings is estimated on the *register-outcome* sample of roughly 705 cohort members who have both age-13 parent-reported items and the longitudinal register outcomes (near-complete for cohort members who did not die or emigrate before the register window opens in 1990). This analysis does not condition on retirement-age survey response: expectation is associated with education (0.290, $p = 0.014$) and preference with mid-career earnings (39.8 thousand kronor, $p = 0.038$).

The subjective well-being analysis (the good-news null) is instead estimated on the smaller *survey-outcome* sample of roughly 340 to 348 members who additionally survived to and responded at 67. The two samples answer different questions under different selection. The paper's primary contrast is carried on the larger, less-selected register sample, whereas the affective nulls rest on the smaller, response-selected survey sample and must be read with the minimum-detectable-effect-size caveat noted in the discussion (the survey models cannot exclude effects below about 0.43 to 0.54 standard deviations of the respective outcome). Reporting register and survey samples side by side complements the complete-case-versus-weighted comparison above.

The compliance items at age 16. Table A14 reports an exploratory analysis on the small subsample of cohort members who, at age 16, reported whether they wanted what their mother or father wanted. In this subsample, preference is strongly associated with attainment (a main effect of 0.717 education levels in the mother variant, $p < 0.01$; 0.477 in the father variant, $p = 0.033$), but the preference-by-compliance interaction is negative (-0.516, $p = 0.060$, mother variant; -0.307, $p = 0.277$, father variant). The pattern (the preference–attainment association is concentrated among children who did *not* say they wanted what their mother wanted) is intriguing but rests on roughly 240 observations and a marginally significant interaction. It is flagged as exploratory and is not part of the paper's claim.

The corresponding well-being interactions are not estimable in this subsample: after the compliance items and retirement-age response are intersected, the cells are too small (about 105–114 members) and contain leverage-one observations that make the HC3 covariance estimator degenerate; those models are therefore not reported.

Multiple testing. Across the exploratory tables, nominal p-values are not adjusted within each small table, but they are read against the cohort-wide multiplicity caution of Section 4.3: only the headline preference-versus-expectation contrast is pre-specified, and the single marginally significant well-being coefficient (realized preference on HAD depression) would not survive a stringent family-wise correction. No exploratory result is hidden, and the one coefficient that changes interpretation under over-control (expectation on education, under post-exposure grade conditioning) is discussed explicitly above.

7 Discussion

Parental push as the organizing interpretation. Read as proxies for parental investment, the estimates describe how a family's push toward a schooling orientation (its direction and its perceived reach) travels into the child's adult life. A parent who expects the academic route invests in getting the child there, and it is the child's realized *schooling* that carries this. A parent who wants that route, whether or not the system is expected to deliver it, may invest in the child in ways that surface later in *earnings* beyond the schooling actually completed.

The clearest sign that these answers register a real, followed direction rather than wishful reporting is the category-level map of Section 5.3: the parent's concrete post-compulsory expectation at 16 lines up with the schooling the child actually reached, from a 7 percent post-secondary share where parents expected immediate work to 75 percent where they expected the long gymnasium.

What the design cannot do is establish that the push *caused* these outcomes. An expectation that tracks later schooling may reflect parental investment that shaped it, or parental foresight that read an ability the school system would have sorted on regardless. The evidence bearing on this (associations that survive conditioning on the child's cognition, and an expectation–schooling link that attenuates when post-exposure school performance is added) is consistent with an investment role without isolating it, and the interpretation below is offered in that spirit.

Interpretation against the framework. The estimates support the H1/H2 ordering. On the schooling margin, expectation dominates: net of preference, family background, and cognition, expectation is associated with higher schooling (0.290 of a level), while preference adds less (0.203, not significant at the 5 percent level). On the earnings margin, preference dominates: 39.8 thousand kronor against expectation's 17.3 (not significant).

Crucially, the pattern respects the human-capital model rather than contradicting it.

Expectation's earnings association is positive and of the magnitude its schooling association would transmit through the Mincer margin. It collapses to 4.3 thousand kronor once completed schooling is held fixed (Section 6): expectation reaches earnings *through* schooling, exactly as equations (3)–(4) imply. Preference retains a positive beyond-schooling component (28.8, $p = 0.110$), consistent with the ν-driven investment channel into uncertified skills, though estimated imprecisely.

Three qualifications belong here. First, the naive ladder shows that *unconditionally* the two measures are nearly interchangeable; the outcome-specific split is a conditional pattern, emerging once family background and cognition absorb the shared information core.

Second, preference's weakness on schooling is sharpest on the six-level scale (ordered logit $p = 0.081$). On the binary post-secondary margin, the probit places preference's marginal effect near expectation's at marginal significance, so H1's ordering reads as "expectation leads," not "preference is absent," on that margin.

Third, the expectation-minus-preference gap (Table 3, Panel C), the direct test of coefficient equality, changes sign across the two margins in the direction H1 and H2 jointly imply but is not individually significant on any single outcome. The two coefficients are ordered as the framework requires and are each significant on its own margin, but they are not far enough apart for the equality test to reject at this sample size. It is thus a robust pattern across specifications, not a sharp gap on a single margin.

The category-level map of Section 5.3 complements the binary contrast: parents' concrete expectation at 16 orders children's realized education levels monotonically, which is what a measure that aggregates selection-relevant information should do.

Magnitude. The expectation's association with attainment (about a third of an education level and 13 percentage points of post-secondary attainment) is sizeable for a single binary parental item measured net of family education and same-wave cognition. The preference's association with mid-career earnings (roughly 40 thousand kronor per year) is likewise meaningful against a cohort mean of about 251 thousand kronor (Table A1). A formal comparison of these magnitudes with the parental-expectations and educational-preferences literature (Sewell et al., 1969) is deferred to a later draft.

The good-news null. The absence of a late-life subjective scar from an unmet academic preference is reassuring, but it should be read with care. Well-being outcomes are observed only for the roughly 590–610 cohort members who survived to and responded at 67; the unmet- and realized-preference models rest on about 340–348. With samples of that size, the analysis is better powered to detect large affective gaps than to rule out small ones.

A minimum-detectable-effect-size calculation makes this concrete: at 80 percent power and the 5 percent level (two-sided), the realized- or unmet-preference contrast in these models can detect effects no smaller than about 0.43 standard deviations of the Satisfaction-with-Life Scale, 0.46 of the single-item life evaluation, 0.48 of loneliness, and 0.54 of depressive symptoms; smaller affective gaps cannot be ruled out at the available retirement-age sample sizes.

External validity. The cohort is a complete school-year cohort in a single midsize Swedish town, observed under a particular educational regime and a particular sequence of macroeconomic episodes. The estimates describe conditional associations in this population. Whether the same preference-versus-expectation contrast would appear in other cohorts, other educational systems, or other labor markets is an open question that this design cannot settle.

Limitations. Three limitations bear on the contribution. First, the design is descriptive: parent-reported measures are not randomly assigned, and estimates do not identify a causal effect of what parents hope for or expect. Second, the parent-questionnaire sample is selected on parental response, and the well-being sample is further selected on survival and response at 67; the response-weighting check addresses selection on observables but cannot address selection on unobservables. Third, the two measures are single binary items collected at one age; richer multi-item or repeated measurement would sharpen the outcome-specific split of expectation from preference.

8 Conclusion

In a complete school-year cohort followed for more than half a century, two parent-reported measures recorded in the same adolescent questionnaire (what parents *preferred* for their 13-year-old and what they *expected* the child would do) serve as proxies for how a family pushed the child toward a schooling orientation. Net of family background and the child's measured cognition, expectation leads on the child's later schooling and reaches earnings through schooling, while preference carries the child's later earnings partly beyond schooling. An unmet academic preference leaves no measurable subjective scar at retirement age.

That these answers register a real, followed direction rather than idle hope is clearest at the category level: parents' concrete post-compulsory expectations at 16 order children's realized education levels monotonically, from a 7 percent post-secondary share where parents expected immediate work to 75 percent where they expected the long gymnasium.

What the design cannot do is separate parental *push* that may have shaped these outcomes from parental *foresight* that may merely have read them; the associations are consistent with an investment channel without isolating it. The natural next step is richer measurement (multi-item or repeated elicitation of the two measures, and a design with exogenous variation in parental investment) to move from a documented association toward an identified effect.

References

Benjamini, Yoav and Yosef Hochberg (1995) "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society: Series B*, 57 (1), 289–300, 10.1111/j.2517-6161.1995.tb02031.x.

Bergman, Lars R., Anna-Karin Andershed, Anna Meehan, and Henrik Andershed (2018) "Individual Development and Adaptation (IDA): A life-span longitudinal program suited for person-oriented research," *Journal for Person-Oriented Research*, 4 (2), 63–77, 10.17505/jpor.2018.07.

Böhlmark, Anders and Matthew J. Lindquist (2006) "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden," *Journal of Labor Economics*, 24 (4), 879–896, 10.1086/506489.

Clark, Andrew E., Ed Diener, Yannis Georgellis, and Richard E. Lucas (2008) "Lags and leads in life satisfaction: A test of the baseline hypothesis," *The Economic Journal*, 118 (529), F222–F243, 10.1111/j.1468-0297.2008.02150.x.

Cunha, Flavio and James J. Heckman (2007) "The technology of skill formation," *American Economic Review*, 97 (2), 31–47, 10.1257/aer.97.2.31.

——— (2008) "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation," *Journal of Human Resources*, 43 (4), 738–782, 10.3368/jhr.43.4.738.

Diener, Ed, Robert A. Emmons, Randy J. Larsen, and Sharon Griffin (1985) "The Satisfaction with Life Scale," *Journal of Personality Assessment*, 49 (1), 71–75, 10.1207/s15327752jpa4901_13.

Englund, Peter (1999) "The Swedish banking crisis: roots and consequences," *Oxford Review of Economic Policy*, 15 (3), 80–97, 10.1093/oxrep/15.3.80.

Esping-Andersen, Gøsta (1990) *The Three Worlds of Welfare Capitalism*, Cambridge: Polity Press, https://press.princeton.edu/books/paperback/9780691028576/the-three-worlds-of-welfare-capitalism.

Haider, Steven and Gary Solon (2006) "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, 96 (4), 1308–1320, 10.1257/aer.96.4.1308.

Heckman, James J., Jora Stixrud, and Sergio Urzua (2006) "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor Economics*, 24 (3), 411–482, 10.1086/504455.

Holmlund, Helena (2007) "A researcher's guide to the Swedish compulsory school reform," Working Paper Series 9/2007, Swedish Institute for Social Research, Stockholm University, https://ideas.repec.org/p/hhs/ifauwp/2020_014.html, Re-released as IFAU Working Paper 2020:15.

Horvitz, Daniel G. and Donovan J. Thompson (1952) "A generalization of sampling without replacement from a finite universe," *Journal of the American Statistical Association*, 47 (260), 663–685, 10.1080/01621459.1952.10483446.

Layard, Richard, Andrew E. Clark, Francesca Cornaglia, Nattavudh Powdthavee, and James Vernoit (2014) "What predicts a successful life? A life-course model of well-being," *The Economic Journal*, 124 (580), F720–F738, 10.1111/ecoj.12170.

Lundborg, Petter, Anton Nilsson, and Dan-Olof Rooth (2014) "Parental education and offspring outcomes: Evidence from the Swedish compulsory school reform," *American Economic Journal: Applied Economics*, 6 (1), 253–278, 10.1257/app.6.1.253.

MacKinnon, James G. and Halbert White (1985) "Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties," *Journal of Econometrics*, 29 (3), 305–325, 10.1016/0304-4076(85)90158-7.

Magnusson, David (1988) *Individual Development from an Interactional Perspective: A Longitudinal Study*, Hillsdale, NJ: Lawrence Erlbaum, https://www.routledge.com/Individual-Development-from-an-Interactional-Perspective-A-Longitudinal-Study/Magnusson/p/book/9781138854222.

Meghir, Costas and Mårten Palme (2005) "Educational reform, ability, and family background," *American Economic Review*, 95 (1), 414–424, 10.1257/0002828053828671.

Pavot, William and Ed Diener (1993) "Review of the Satisfaction with Life Scale," *Psychological Assessment*, 5 (2), 164–172, 10.1037/1040-3590.5.2.164.

Schwandt, Hannes (2016) "Unmet aspirations as an explanation for the age U-shape in wellbeing," *Journal of Economic Behavior and Organization*, 122, 75–87, 10.1016/j.jebo.2015.11.011.

Sewell, William H., Archibald O. Haller, and Alejandro Portes (1969) "The educational and early occupational attainment process," *American Sociological Review*, 34 (1), 82–92, 10.2307/2092789.

Wooldridge, Jeffrey M. (2007) "Inverse probability weighted estimation for general missing data problems," *Journal of Econometrics*, 141 (2), 1281–1301, 10.1016/j.jeconom.2007.02.002.

A Supplementary tables and figures

Table A1. Descriptive statistics by sex

Variable	All	Women	Men	p	N	Missing
Parent-reported preference: academic orientation, 13 (=1)	0.71 (0.45)	0.73 (0.44)	0.69 (0.46)	0.147	1030	362
Parent-reported expectation: academic orientation, 13 (=1)	0.63 (0.48)	0.65 (0.48)	0.61 (0.49)	0.133	1028	364
Preference only: academic preference, vocational expectation, 13 (=1)	0.09 (0.29)	0.09 (0.29)	0.09 (0.29)	0.929	1015	377
Academic education suits the child, 13 (=1)	0.60 (0.49)	0.63 (0.48)	0.57 (0.50)	0.042**	981	411
Academic education pays off today, 13 (=1)	0.77 (0.42)	0.77 (0.42)	0.77 (0.42)	0.860	897	495
Parent discussed orientation with child, 13 (=1)	0.98 (0.13)	0.98 (0.15)	0.99 (0.11)	0.155	1063	329
Discuss the future with child, 16 (1-4, high=more)	2.92 (0.67)	2.88 (0.68)	2.96 (0.65)	0.077*	935	457
Suited for a long academic education, 16 (1-5, high=suited)	3.40 (1.15)	3.40 (1.13)	3.40 (1.17)	0.944	878	514
Parent and pupil agree on school type, 16 (=1)	0.87 (0.33)	0.87 (0.34)	0.88 (0.32)	0.512	896	496
Wants what mother wants, 16 (=1)	0.67 (0.47)	0.76 (0.43)	0.56 (0.50)	0.000***	437	955
Wants what father wants, 16 (=1)	0.68 (0.47)	0.75 (0.43)	0.62 (0.49)	0.006***	421	971
Highest education level 1990–2022 (SUN 1–6)	3.57 (1.11)	3.69 (1.13)	3.46 (1.08)	0.000***	1344	48
Post-secondary attainment (SUN>=4) (=1)	0.39 (0.49)	0.43 (0.50)	0.35 (0.48)	0.004***	1344	48
Parent's academic preference UNMET (=1)	0.34 (0.47)	0.33 (0.47)	0.35 (0.48)	0.459	992	400
Parent's academic preference REALIZED (=1)	0.37 (0.48)	0.40 (0.49)	0.33 (0.47)	0.031**	992	400
Mean wage ages 45–54 (1,000 SEK)	250.85 (225.92)	209.90 (123.00)	290.47 (287.50)	0.000***	1324	68
Employment share 1990–2022 (0–1)	0.80 (0.24)	0.81 (0.23)	0.80 (0.25)	0.700	1348	44
SWLS at 67 (5–35)	25.91 (6.13)	25.65 (6.32)	26.23 (5.87)	0.244	606	786
OECD-SWL at 67 (0–10)	7.64 (2.12)	7.50 (2.19)	7.81 (2.02)	0.081*	593	799
Loneliness at 67 (UCLA-3, 3-9)	3.92 (1.31)	3.95 (1.33)	3.88 (1.28)	0.426	592	800
HAD depression at 67 (0-21)	2.25 (2.08)	2.17 (2.02)	2.36 (2.16)	0.315	604	788
Family socioeconomic standing, 1970 (1–7, high=high)	2.14 (1.41)	2.19 (1.45)	2.08 (1.37)	0.189	1209	183
Family education, grade 3 (1–7, high=high)	2.92 (1.60)	2.97 (1.61)	2.87 (1.59)	0.312	972	420
Intact family at grade 3 (=1)	0.89 (0.31)	0.90 (0.31)	0.89 (0.31)	0.850	985	407
DBA total intelligence, 13 (six stanines, 6–54)	30.27 (8.98)	30.37 (9.08)	30.16 (8.88)	0.696	1102	290

Note: Cell entries are mean (standard deviation) for continuous variables and proportion (standard deviation) for binary variables (=1). The p column tests the women-versus-men difference (Welch t-test for continuous variables; χ^2 or Fisher exact for binary; Mann–Whitney for skewed variables), with statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***). N is the count of cohort members for whom the variable is observed; Missing is the complement out of the 1,392 cumulative cohort members. Wage is in thousands of Swedish kronor. Scale orientations: every scale is reported so that higher = more of the named construct, except loneliness (UCLA-3) and depression (HAD-D), which run higher = worse. The family socioeconomic standing scale is the family's 1970 census classification with codes outside the documented 1–7 frame set to missing, re-oriented ($8 - x$; raw coding runs 1 = highest); family education (1965 parent questionnaire) and the age-16 suitability and discussion items are re-oriented likewise (raw codings run high = low). The models use the two socioeconomic-group indicators, not the 1–7 scale (Section 3.3).

Table A2. Descriptive statistics by birth-year group

Variable	All	≤1954	1955	≥1956	p	N	Missing
Parent-reported preference: academic orientation, 13 (=1)	0.71 (0.45)	0.44 (0.50)	0.73 (0.44)	1.00 (0.00)	0.000***	1030	362
Parent-reported expectation: academic orientation, 13 (=1)	0.63 (0.48)	0.36 (0.48)	0.65 (0.48)	1.00 (0.00)	0.000***	1028	364
Preference only: academic preference, vocational expectation, 13 (=1)	0.09 (0.29)	0.12 (0.32)	0.09 (0.29)	0.00 (0.00)	0.360	1015	377
Academic education suits the child, 13 (=1)	0.60 (0.49)	0.32 (0.47)	0.63 (0.48)	1.00 (0.00)	0.000***	981	411
Academic education pays off today, 13 (=1)	0.77 (0.42)	0.66 (0.48)	0.78 (0.42)	1.00 (0.00)	0.011**	897	495
Parent discussed orientation with child, 13 (=1)	0.98 (0.13)	0.95 (0.22)	0.99 (0.12)	1.00 (0.00)	0.018**	1063	329
Discuss the future with child, 16 (1-4, high=more)	2.92 (0.67)	2.92 (0.63)	2.92 (0.67)	3.07 (0.73)	0.704	935	457
Suited for a long academic education, 16 (1-5, high=suited)	3.40 (1.15)	3.24 (1.05)	3.40 (1.15)	3.92 (1.26)	0.129	878	514
Parent and pupil agree on school type, 16 (=1)	0.87 (0.33)	0.80 (0.40)	0.88 (0.33)	0.93 (0.27)	0.147	896	496
Wants what mother wants, 16 (=1)	0.67 (0.47)	0.68 (0.47)	0.67 (0.47)	0.60 (0.52)	0.879	437	955
Wants what father wants, 16 (=1)	0.68 (0.47)	0.62 (0.49)	0.69 (0.46)	0.67 (0.50)	0.625	421	971
Highest education level 1990–2022 (SUN 1–6)	3.57 (1.11)	3.18 (1.05)	3.61 (1.11)	3.88 (1.05)	0.000***	1344	48
Post-secondary attainment (SUN>=4) (=1)	0.39 (0.49)	0.23 (0.42)	0.41 (0.49)	0.53 (0.51)	0.000***	1344	48
Parent's academic preference UNMET (=1)	0.34 (0.47)	0.28 (0.45)	0.34 (0.48)	0.50 (0.52)	0.225	992	400
Parent's academic preference REALIZED (=1)	0.37 (0.48)	0.14 (0.35)	0.39 (0.49)	0.50 (0.52)	0.000***	992	400
Mean wage ages 45–54 (1,000 SEK)	250.85 (225.92)	176.02 (145.22)	258.81 (232.96)	299.21 (123.99)	0.000***	1324	68
Employment share 1990–2022 (0–1)	0.80 (0.24)	0.71 (0.30)	0.81 (0.24)	0.83 (0.21)	0.000***	1348	44
SWLS at 67 (5–35)	25.91 (6.13)	24.08 (6.18)	26.08 (6.13)	25.30 (4.52)	0.092*	606	786
OECD-SWL at 67 (0–10)	7.64 (2.12)	7.13 (2.41)	7.69 (2.08)	7.00 (2.58)	0.150	593	799
Loneliness at 67 (UCLA-3, 3-9)	3.92 (1.31)	4.04 (1.24)	3.91 (1.32)	3.90 (1.20)	0.633	592	800
HAD depression at 67 (0-21)	2.25 (2.08)	3.08 (2.30)	2.19 (2.06)	1.60 (1.51)	0.007***	604	788
Family socioeconomic standing, 1970 (1–7, high=high)	2.14 (1.41)	1.85 (1.09)	2.16 (1.43)	2.47 (1.66)	0.050*	1209	183
Family education, grade 3 (1–7, high=high)	2.92 (1.60)	2.36 (1.43)	2.96 (1.60)	4.00 (1.65)	0.000***	972	420
Intact family at grade 3 (=1)	0.89 (0.31)	0.87 (0.34)	0.89 (0.31)	1.00 (0.00)	0.358	985	407
DBA total intelligence, 13 (six stanines, 6–54)	30.27 (8.98)	23.76 (7.42)	30.81 (8.83)	38.53 (8.10)	0.000***	1102	290

Note: As Table A1 (including its scale-orientation conventions), with columns for the three birth-year groups (≤1954 / 1955 / ≥1956) and a k-group test (F-test, χ^2, or Kruskal–Wallis as appropriate). The age-13 items are degenerate in the small ≥1956 group; birth year is used in the models only as a three-level control.

Table A3. Descriptive statistics by cohort-entry route

Variable	All	Original-W1	Incomers	Movers	p	N	Missing
Parent-reported preference: academic orientation, 13 (=1)	0.71 (0.45)	0.70 (0.46)	0.77 (0.42)		0.124	991	319
Parent-reported expectation: academic orientation, 13 (=1)	0.64 (0.48)	0.63 (0.48)	0.71 (0.46)		0.079*	989	321
Preference only: academic preference, vocational expectation, 13 (=1)	0.09 (0.28)	0.09 (0.29)	0.07 (0.26)		0.527	977	333
Academic education suits the child, 13 (=1)	0.61 (0.49)	0.60 (0.49)	0.66 (0.47)		0.183	942	368
Academic education pays off today, 13 (=1)	0.77 (0.42)	0.77 (0.42)	0.79 (0.41)		0.572	860	450
Parent discussed orientation with child, 13 (=1)	0.98 (0.13)	0.98 (0.14)	0.99 (0.09)		0.383	1022	288
Discuss the future with child, 16 (1-4, high=more)	2.92 (0.67)	2.91 (0.65)	2.96 (0.70)		0.357	891	419
Suited for a long academic education, 16 (1-5, high=suited)	3.42 (1.14)	3.39 (1.16)	3.54 (1.06)		0.106	838	472
Parent and pupil agree on school type, 16 (=1)	0.88 (0.33)	0.88 (0.33)	0.87 (0.33)		0.792	860	450
Wants what mother wants, 16 (=1)	0.68 (0.47)	0.66 (0.47)	0.74 (0.44)		0.181	406	904
Wants what father wants, 16 (=1)	0.69 (0.46)	0.67 (0.47)	0.76 (0.43)		0.129	392	918
Highest education level 1990–2022 (SUN 1–6)	3.60 (1.11)	3.59 (1.12)	3.68 (1.08)	3.14 (1.15)	0.080*	1267	43
Post-secondary attainment (SUN>=4) (=1)	0.40 (0.49)	0.39 (0.49)	0.43 (0.50)	0.29 (0.46)	0.261	1267	43
Parent's academic preference UNMET (=1)	0.34 (0.47)	0.34 (0.47)	0.35 (0.48)		0.724	956	354
Parent's academic preference REALIZED (=1)	0.37 (0.48)	0.37 (0.48)	0.41 (0.49)		0.407	956	354
Mean wage ages 45–54 (1,000 SEK)	254.56 (228.84)	254.10 (205.45)	264.33 (303.62)	153.69 (116.52)	0.102	1249	61
Employment share 1990–2022 (0–1)	0.81 (0.24)	0.82 (0.23)	0.79 (0.25)	0.66 (0.35)	0.006***	1271	39
SWLS at 67 (5–35)	25.97 (6.05)	26.03 (6.09)	25.81 (5.97)	24.86 (5.27)	0.831	593	717
OECD-SWL at 67 (0–10)	7.65 (2.10)	7.64 (2.16)	7.65 (1.90)	8.14 (1.77)	0.821	580	730
Loneliness at 67 (UCLA-3, 3-9)	3.90 (1.30)	3.88 (1.27)	4.04 (1.40)	3.29 (0.49)	0.209	578	732
HAD depression at 67 (0-21)	2.19 (2.02)	2.15 (1.99)	2.33 (2.10)	2.14 (2.48)	0.655	589	721
Family socioeconomic standing, 1970 (1–7, high=high)	2.14 (1.40)	2.05 (1.25)	2.47 (1.81)	2.06 (1.53)	0.000***	1147	163
Family education, grade 3 (1–7, high=high)	2.92 (1.60)	2.92 (1.59)		2.68 (1.95)	0.523	972	338
Intact family at grade 3 (=1)	0.89 (0.31)	0.90 (0.30)		0.70 (0.47)	0.005***	985	325
DBA total intelligence, 13 (six stanines, 6–54)	30.43 (8.89)	30.35 (8.67)	30.99 (10.28)		0.440	1059	251

Note: As Table A1, by cohort entry (original grade-three members, later incomers, movers). Family-structure items are largely unobserved for incomers, who joined after the grade-three family measurement; the corresponding cells are blank.

Table A4. Binary parental and outcome items: contrasts across sex, birth-year group, and cohort entry

	Sex			Birth-year			Cohort entry		
Variable	Women	Men	p	1955	≤1954	p	Orig.-W1	Incomers	p
Parent-reported preference: academic orientation, 13 (=1)	0.729	0.688	0.147	0.732	0.442	0.000***	0.703	0.770	0.355
Parent-reported expectation: academic orientation, 13 (=1)	0.654	0.609	0.133	0.654	0.358	0.000***	0.626	0.707	0.402
Preference only: academic preference, vocational expectation, 13 (=1)	0.090	0.091	0.929	0.089	0.116	0.668	0.091	0.074	0.880
Academic education suits the child, 13 (=1)	0.634	0.571	0.042**	0.626	0.323	0.000***	0.599	0.664	0.639
Academic education pays off today, 13 (=1)	0.768	0.773	0.860	0.778	0.658	0.106	0.769	0.794	0.840
Parent discussed orientation with child, 13 (=1)	0.977	0.989	0.155	0.986	0.948	0.031**	0.981	0.992	0.337
Parent and pupil agree on school type, 16 (=1)	0.867	0.881	0.512	0.879	0.803	0.142	0.879	0.872	0.383
Wants what mother wants, 16 (=1)	0.759	0.561	0.000***	0.671	0.684	0.952	0.664	0.739	0.613
Wants what father wants, 16 (=1)	0.750	0.624	0.006***	0.692	0.619	0.353	0.670	0.756	0.308
Post-secondary attainment (SUN>=4) (=1)	0.427	0.351	0.004***	0.406	0.225	0.000***	0.392	0.434	0.642
Parent's academic preference unmet (=1)	0.329	0.351	0.459	0.344	0.278	0.420	0.337	0.353	0.660
Parent's academic preference realized (=1)	0.399	0.333	0.031**	0.387	0.144	0.000***	0.365	0.405	0.945
Intact family at grade 3 (=1)	0.895	0.892	0.850	0.895	0.868	0.731	0.897		0.014**

Note: Proportions for the binary items in each group, with the difference test for each pairing (χ^2 or Fisher exact). Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***). A blank entry indicates a cell with insufficient observations. For legibility the birth-year and cohort-entry panels show only the two largest groups in each; the full three-group descriptives are in Appendix Tables A2 and A3.

Table A5. Headline attainment models with the parent-reported preference and expectation: full coefficient set

Variable	Educ. level	Post-sec. (=1)	Empl. share	Wage 45–54 (kSEK)
Parent's expectation: academic orientation (=1)	0.290** (0.118)	0.126** (0.056)	-0.001 (0.029)	17.291 (18.747)
Parent's preference: academic orientation (=1)	0.203* (0.122)	0.088 (0.057)	0.053* (0.032)	39.782** (19.179)
Woman (=1)	0.241*** (0.072)	0.077** (0.033)	0.013 (0.016)	-89.636*** (16.304)
Born 1954 or earlier (=1)	0.077 (0.124)	0.009 (0.055)	-0.002 (0.035)	-39.391** (19.453)
Born 1956 or later (=1)	-0.648* (0.336)	-0.292 (0.191)	-0.045 (0.066)	-28.886 (56.275)
Family socioeconomic group: white-collar (=1)	-0.030 (0.099)	-0.006 (0.046)	0.003 (0.022)	-4.393 (19.626)
Family socioeconomic group: self-employed/farm (=1)	0.399* (0.240)	0.282*** (0.090)	-0.103 (0.091)	-70.216 (49.872)
Family education (1–7, high = high)	0.157*** (0.024)	0.066*** (0.011)	0.011** (0.006)	17.819** (7.883)
Intact family (=1)	-0.016 (0.153)	0.008 (0.068)	0.018 (0.035)	-20.552 (25.469)
Cognition at 13 (six sub-test stanines, 6–54)	0.035*** (0.005)	0.015*** (0.002)	0.002* (0.001)	3.659*** (0.998)
Observations	705	705	707	699

Note: Full coefficient set for the C1 attainment and labor-market models of Table 3, all four register outcomes; heteroskedasticity-consistent (HC3) standard errors in parentheses. Same sample and specification as Table 3; this table restores the control coefficients suppressed there. Reference categories: born 1955; worker-group family. Family education and the cognition score are re-oriented and summed so that higher = more (Section 3.3). Statistical significance at $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table A6. Preference-by-expectation group profile (exploratory)

	Both academic	Both vocational	Preference only	Expectation only
N	628	282	92	13
Family socioeconomic standing, 1970 (1–7, high = high)	2.10	1.92	2.16	2.36
Family education, grade 3 (1–7, high = high)	3.34	2.00	2.51	2.75
Cognition at 13 (six stanines, 6–54)	33.88	24.13	27.87	25.85
Share women	0.52	0.47	0.50	0.46
Education level 1990–2022 (SUN 1–6)	3.97	2.94	3.39	3.08
Post-secondary (SUN $\geq$ 4)	0.56	0.12	0.28	0.08
Mean wage ages 45–54 (1,000 SEK)	296.72	180.98	221.11	176.19
SWLS at 67 (5–35)	26.37	25.19	25.26	25.33

Note: Group means for the four groups formed by crossing the age-13 academic preference and academic expectation: *both academic* (preferred and expected the academic orientation), *both vocational* (preferred and expected the vocational orientation), *preference only* (preferred the academic orientation but expected the vocational one), and *expectation only* (expected the academic orientation but did not prefer it). Cognition is the summed-stanine intelligence score at age 13; education is the highest level 1990–2022; wage is mean annual wage income over ages 45–54 (1,000 SEK); SWLS is the Satisfaction With Life Scale at 67 (higher = more satisfied). The socioeconomic-standing and family-education scales are re-oriented so that higher = higher; see the note to Appendix Table A1. Exploratory; cell sizes vary across rows because of differential missingness.

Table A7. Preference-by-expectation groups on adult register outcomes (exploratory)

	Education (SUN)	Post-sec. (LPM)	Wage 45–54	Emp. share
Both academic	0.482*** (0.101)	0.205*** (0.044)	55.050*** (14.898)	0.052** (0.024)
Preference only	0.129 (0.139)	0.028 (0.062)	26.078 (20.655)	0.049 (0.035)
Expectation only	-0.054 (0.220)	-0.151 (0.108)	-45.553 (47.418)	-0.022 (0.087)
N	705	705	699	707

Note: Ordinary least squares estimates with HC3 robust standard errors in parentheses. The rows are the four expectation-by-preference groups defined in Appendix Table A6; the omitted reference group is the *both-vocational* families (preferred and expected the vocational orientation). "Both academic" both preferred and expected the academic orientation; "preference only" preferred it but expected the vocational orientation; "expectation only" expected it but did not prefer it. Outcomes and controls as in Table 3. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***). Exploratory analysis; the preference-only and expectation-only cells are small (92 and 13 members).

Table A8. Unmet versus realized academic preference on subjective well-being at 67 (secondary, exploratory)

	SWLS	OECD-SWL	UCLA-3	HAD-D
Academic preference unmet	0.898 (0.975)	0.023 (0.366)	0.145 (0.228)	-0.475 (0.397)
Academic preference realized	0.946 (0.941)	0.023 (0.347)	-0.039 (0.223)	-0.734* (0.398)
N	348	340	342	344

Note: Ordinary least squares estimates with HC3 robust standard errors in parentheses; the reference category is no academic preference. SWLS is the Satisfaction With Life Scale (5–35, higher = more satisfied); OECD-SWL is the single-item life-satisfaction measure (0–10, higher = more satisfied); UCLA-3 is loneliness (3–9, higher = lonelier); HAD-D is the depression subscale (0–21, higher = more depressive symptoms). Controls as in Table 3. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***). Secondary, exploratory analysis (Section 5.5). A response-weighted version is in Appendix Table A13.

Table A9. Attainment models conditioning on post-exposure school grades at 16

	Education (SUN)	Post-sec. (LPM)
Parent's preference: academic orientation	0.175 (0.123)	0.086 (0.058)
Parent's expectation: academic orientation	0.187 (0.120)	0.079 (0.057)
School grades at 16 (mean of 1–5 scale, high = better)	0.569*** (0.056)	0.235*** (0.026)
N	657	657

Note: As the first two columns of Table 3, adding the child's mean school grades at age 16 as a covariate (mean of marks on the 1–5 scale, higher = better performance). School grades are post-exposure (measured after the age-13 items) and absorb information in the measures that operates through realized school performance; this is a conservative over-adjustment, not the headline specification. HC3 robust standard errors in parentheses. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A10. Ordered- and binary-response confirmation of the attainment results

	Ordered logit: education (SUN 1–6)		Probit: post-secondary	
	Coefficient	AME P(SUN≥4)	Coefficient	AME
Parent's preference: academic orientation	0.477* (0.274)	0.092	0.376* (0.196)	0.113
Parent's expectation: academic orientation	0.598** (0.256)	0.118	0.340** (0.172)	0.102
N	705		705	

Note: Same sample ($N = 705$) and right-hand side as the education columns of Table 3. The ordered logit treats the six-level SUN education scale as ordered; coefficients are in log-odds with maximum-likelihood standard errors in parentheses, and the AME column reports the average marginal effect on the probability of reaching at least the post-secondary level (SUN $\geq$ 4), computed by predicting each observation with the indicator set to one versus zero and averaging the difference. The probit models the post-secondary indicator; coefficients carry HC1 heteroskedasticity-robust standard errors and the AME column reports average marginal effects (HC3 is defined for the least-squares models only). Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A11. Earnings conditional on completed schooling (the transmission check of H2)

	Mean wage 45–54 (1,000 SEK)
Parent's preference: academic orientation	28.815 (18.046)
Parent's expectation: academic orientation	4.266 (17.597)
Education level (SUN 1–6)	48.300*** (7.620)
N	697

Note: Ordinary least squares with HC3 robust standard errors in parentheses. As the wage column of Table 3, adding the completed education level (SUN 1–6, higher = more education) as a covariate; schooling is post-exposure relative to the age-13 parent-reported items, so this is a mechanism decomposition in the sense of equation (3), not a headline specification. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A12. Person-year panel estimates with year fixed effects

	Annual wage, ages 45–54 (1,000 SEK)	Employed in year t (0/1, 1990–2022)
Parent's preference: academic orientation	38.551** (19.004)	0.026* (0.015)
Parent's expectation: academic orientation	17.422 (18.509)	0.000 (0.013)
Year fixed effects	Yes	Yes
Person-years	6901	49187
Persons	699	728

Note: Ordinary least squares estimates of equation (16) on the person-year register panel (the employment column is a linear probability model); year fixed effects included; standard errors clustered on the person in parentheses. The wage column uses person-years at ages 45–54 with observed wage income (in thousands of Swedish kronor); the employment column uses all register person-years 1990–2022. Controls as in Table 3. Person fixed effects cannot be used because the parent-reported measures and the controls are time-invariant. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A13. Subjective well-being at 67, response-weighted

	SWLS	OECD-SWL	UCLA-3	HAD-D
Academic preference unmet	0.903 (1.144)	0.112 (0.451)	0.136 (0.321)	-0.551 (0.502)
Academic preference realized	1.152 (1.109)	0.156 (0.439)	-0.195 (0.299)	-0.928* (0.497)
N	348	340	342	344

Note: As Table A8 (including its scale ranges and orientations), estimated by weighted least squares with inverse-probability-of-response weights for the retirement-age wave. HC3 robust standard errors in parentheses. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A14. Age-16 compliance interaction on educational attainment (exploratory)

	Education (SUN), age-16 compliance subsample	
	Mother variant	Father variant
Parent's preference: academic orientation	0.717*** (0.210)	0.477** (0.224)
Child wants what parent wants, 16	0.337 (0.219)	0.352 (0.230)
Preference × compliance	-0.516* (0.274)	-0.307 (0.282)
N	244	240

Note: Ordinary least squares estimates with HC3 robust standard errors in parentheses, on the subsample of cohort members who reported at age 16 whether they wanted what their mother (or father) wanted. "Preference × compliance" is the interaction of the parent's academic preference with the child's reported compliance. Controls as in Table 3. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***). Exploratory and small-sample; the well-being variants are not estimable (see Section 6).

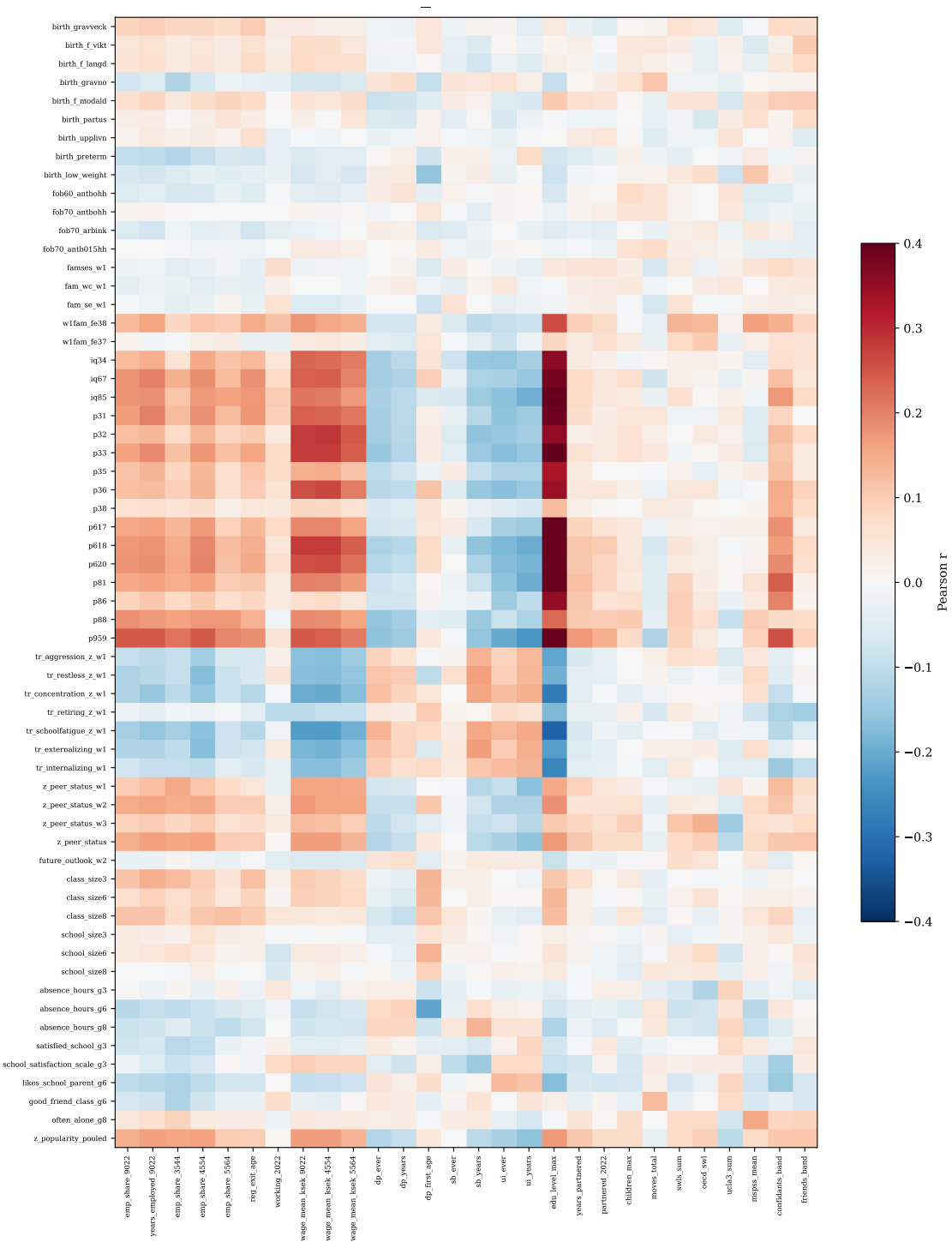

Figure A1. Cohort-wide battery-association screen

Note: Heatmap of the cohort-wide association screen of candidate measures against the register outcomes, computed on the full cohort (N = 1,392). The screen is a descriptive multiplicity check (1,708 tests; 540 significant at a Benjamini–Hochberg false-discovery rate below 0.05) and motivates the multiplicity caution applied to the exploratory analyses; it is not part of the confirmatory contrast.

Table A15. Joint cross-margin test of the preference and expectation outcome-specific split

	Schooling (SD)	Earnings (SD)
Expectation−preference gap (β_G)	+0.029	-0.050
	(0.099)	(0.081)
Cross-margin flip $\beta_G^{\text{earn}} - \beta_G^{\text{sch}} = -0.078$ (se 0.102, $p = 0.441$)		
Joint Wald ($\beta_G = 0$ on both) $\chi^2_2 = 0.70$, $p = 0.706$		
N (persons)	697	

Note: The two headline outcomes (education level and mean wage over ages 45–54) are standardized to zero mean and unit variance on the common complete-case sample ($N = 697$) and stacked, giving two observations per person. The expectation−preference gap $\beta_G = (\beta_E - \beta_P)/2$ and the academic-orientation sum are each fully interacted with an earnings indicator and estimated by ordinary least squares with standard errors clustered on the person; controls as in Table 3. The cross-margin flip is the difference in the gap coefficient between the earnings and the schooling equation; the joint Wald test is that the gap is zero on both margins. Neither test rejects, so the outcome-specific split is a directional pattern rather than a statistically established difference. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A16. Who answered the 1968 parental questionnaire: responders versus non-responders

	Responders	Non-responders	Std. diff.
Woman (=1)	0.502	0.465	+0.074
Census family SES, 1970 (1–7, higher = higher standing)	2.055	2.397	-0.226***
Incomer (=1)	0.118	0.462	-0.816***
Register education 1990–2022 (SUN 1–6)	3.603	3.502	+0.090
Mean wage ages 45–54 (1,000 SEK)	255.365	243.668	+0.049
Cognition at 13 (6–54)	30.469	27.417	+0.352***

Note: Means for cohort members present at the 1968 (grade-six) wave (the 22 movers are excluded, as the item was structurally undefined for them), by whether the parent answered the schooling-orientation item. "Std. diff." is the difference in means divided by the pooled standard deviation; the p-value is from a Welch two-sample t-test. Responders differ from non-responders on inputs (cohort entry, family socioeconomic standing, and cognition) but not on the adult register outcomes (education and wage). Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A17. Headline estimates with inverse-probability-of-response weights for the 1968 parental questionnaire

	Education (SUN)		Wage 45–54	
	OLS	IPW	OLS	IPW
Parent's preference: academic orientation	0.203* (0.122)	0.210* (0.122)	39.782** (19.179)	40.067** (19.188)
Parent's expectation: academic orientation	0.290** (0.118)	0.284** (0.118)	17.291 (18.747)	16.855 (18.745)
N	705	705	699	699

Note: Ordinary least squares (OLS) and weighted least squares (IPW) estimates of the additive model of Table 3 Panel A for the two headline outcomes. The weights are stabilized inverse probabilities of answering the 1968 parental questionnaire, from a logit response model estimated over the non-mover cohort on sex, birth-year group, cohort-entry status, and 1970 census family socioeconomic standing (characteristics available for the whole cohort); the weights range from 0.84 to 2.55. HC3 robust standard errors in parentheses. Reweighting leaves the estimates essentially unchanged. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

Table A18. The headline earnings result across alternative earnings windows

	Ages 40–49	Ages 45–54 (headline)	Ages 50–59	Ages 55–64	Ages 40–59	All register years
Parent's preference: academic orientation	26.37* (14.92)	39.78** (19.18)	42.98* (22.93)	38.65* (22.84)	35.90** (17.84)	26.19* (13.77)
Parent's expectation: academic orientation	16.11 (14.49)	17.29 (18.75)	17.00 (22.50)	11.32 (22.42)	17.17 (17.57)	16.37 (13.66)
Mean wage (1,000 SEK)	214.6	262.9	303.1	302.7	257.3	209.2
N	702	699	688	681	702	728

Note: Ordinary least squares estimates of the additive model of Table 3 Panel A with the earnings outcome recomputed as the person-level mean annual wage over the age window in each column, built from the annual register data (1990–2022). "Ages 45–54" is the headline window; "All register years" averages every observed year. Controls as in Table 3; HC3 robust standard errors in parentheses. The preference is positive and at least marginally significant on earnings in every window and the expectation is significant in none, so the contrast is not an artifact of the mid-career window. Statistical significance at $p < 0.10$ (*), $p < 0.05$ (**), $p < 0.01$ (***).

B Construction of the analytic file

The analytic file is built from a verified cohort spine of the 1,392 cumulative members and merged, by the cohort's anonymized identifiers, with three derived sources: (i) the parental-expectations items and well-being items from the survey battery, (ii) the same-wave cognition measure, and (iii) the register attainment and labor-market outcomes. "Don't know" codes on the age-16 ordinal items are recoded to missing before use.

Preference and expectation are never imputed. Cohort members missing preference or expectation are dropped from the models that use them, and the realized analytic N is reported in every table. The complete-case sample is primary. A coherence check of the preference item against a separately worded, reverse-coded education-preference item confirms the polarity of the coding.

Scale orientations are aligned at construction so that higher values mean more of the named construct, with each reversal documented in the construction scripts: the family's 1970 census socioeconomic classification (raw coding 1 = highest standing; codes 8–9, outside the documented 1–7 frame, set to missing; re-oriented $x' = 8 - x$), the 1965 family-education scale (raw 1 = highest; re-oriented $x' = 8 - x$), the age-16 suitability item (raw 1 = surely suited; re-oriented $x' = 6 - x$), and the age-16 discussion item (re-oriented $x' = 5 - x$). The depression subscale (HAD-D) is pre-oriented at source (0 = no symptom per item) and is summed without reversal; no already-oriented scale is reversed. All derived data files are stored in a format readable by current versions of standard statistical software.

C Missing-data summary

Counts of observed and missing values for each analytic variable (out of 1,392 cohort members in total) are reported in the N and Missing columns of Tables A1, A2, and A3.

In brief, parent-reported items are observed for about 1,030 members; register outcomes for about 1,324–1,348 of the 1,348 register-eligible members; and well-being outcomes for about 590–610 members who responded at age 67. The primary attainment models use a complete-case sample of 699–707 members. The well-being models use 340–348. The exact N varies slightly by outcome and model specification.

Missingness in the parental items is largely structural (Section 4.1): members who left after grade three or joined after the 1968 wave could not receive the parent questionnaire.