

Skill Complementarity in Educational Production: Evidence from Irish Secondary Students

Beatriz Gietner

University College Dublin

March 14, 2026

Abstract

How noncognitive skills are measured shapes whether they appear as complements or substitutes to cognition in educational production. Using data from the Growing Up in Ireland longitudinal study, I estimate translog production functions for Maths and English achievement among Irish secondary students and show that the estimated degree of skill complementarity depends on the noncognitive instrument. When the noncognitive input is a school-proximal behavioural measure (SDQ Focused Behaviour), elasticity-of-substitution estimates are consistently below one, indicating that cognitive and behavioural skills are hard to trade off against each other. This result is robust to baseline controls, alternative scaling, bootstrap inference, and extension to Leaving Certificate outcomes. When the noncognitive input is instead measured with a broad personality inventory (TIPI Conscientiousness), the same complementarity pattern weakens, precision falls, and one subgroup (girls in Maths) produces a point estimate above unity. The contribution is that measurement choices on the noncognitive margin matter for production-function conclusions: skill complementarity is empirically stable for attentional regulation but not for generic personality proxies.

Keywords: cognitive skills, noncognitive skills, educational production, academic performance, gender differences, human capital

JEL Codes: I21, I24, J24, C31

1 Introduction

A large literature establishes that cognitive and noncognitive skills each predict academic achievement (Heckman et al., 2006; Kautz et al., 2014), but there is far less evidence on

how these inputs combine in production. Such distinction has economic consequences: if skills are close substitutes, progress on one margin can offset weakness on the other, whereas if they are complements, balanced development becomes central and single-input interventions face quickly diminishing returns. The question of substitutability therefore shapes how human-capital theory models the returns to balanced versus specialised skill development.

This paper estimates the degree of substitutability between cognitive and noncognitive skills in producing Maths and English achievement among Irish secondary students. Using data from the Growing Up in Ireland longitudinal study, I estimate linear and translog production functions and compute elasticity-of-substitution (ES) parameters by subject and gender, treating linear specifications that condition on earlier childhood skill measures as the preferred benchmark. Ireland provides a relatively strong empirical setting for this exercise because its centralised curriculum and standardised examinations deliver a consistent achievement metric, while current policy emphasis on student wellbeing keeps the noncognitive margin substantively relevant (Lawlor, 2019; National Council for Curriculum and Assessment, 2021).

The central finding is that the estimated degree of skill complementarity depends on how the noncognitive margin is measured. When the noncognitive input is a school-proximal behavioural indicator - the hyperactivity/inattention dimension of the Strengths and Difficulties Questionnaire (SDQ), recoded as Focused Behaviour - ES estimates are consistently below one across subjects and subgroups, indicating that cognitive and behavioural skills are hard to substitute for one another. This result survives baseline controls, alternative scaling, and bootstrap inference. When broader personality traits from the Ten Item Personality Inventory (TIPI) replace these behavioural indicators, the same complementarity pattern weakens, precision falls, and one subgroup (girls in Maths) produces a point estimate above unity with wide confidence intervals. The contribution is therefore that measurement choices on the noncognitive margin are consequential for production-function conclusions: complementarity is empirically stable for attentional regulation but not for generic personality proxies, a distinction that matters for how researchers specify and interpret skill inputs in education.

To avoid conceptual ambiguity, I distinguish two notions of complementarity that appear in the analysis: negative linear interaction terms capture marginal cross-effects (Edgeworth substitutability), while an $ES < 1$ in the translog captures low substitutability along isoquants (Hicks complementarity). These are different objects and can coexist.

The main identification comes from temporal ordering, cross-student variation in observed inputs, and rich controls, rather than from exogenous treatment assignment. I interpret all parameters as conditional associations for the analysis sample. The results are therefore informative about empirical patterns and targeting priorities, but not sufficient

for causal claims about intervention effects.

The remainder of the chapter is organised as follows. The related literature is reviewed below. Section 2 describes the data and measurement instruments. Section 3 presents the linear production-function framework and results. Section 4 extends the analysis to the translog specification, reports ES estimates and robustness checks, and discusses boundary conditions. Section 5 concludes.

1.1 Related literature

1.1.1 Education production functions

The education production function has been a central organising framework in the economics of education since Hanushek (2003), who formalised the relationship between measurable school and family inputs and student achievement. The large empirical literature that followed has overwhelmingly used linear or log-linear specifications with school-resource inputs (class size, expenditure, teacher experience), finding weak and inconsistent relationships with outcomes (Hanushek & Woessmann, 2012). Todd and Wolpin (2003) clarified the identifying assumptions required for different empirical specifications - contemporaneous, value-added, and cumulative - establishing the methodological standard for the field.

A separate strand has applied flexible functional forms - constant elasticity of substitution (CES), translog - to production function estimation in manufacturing, agriculture, and energy (Christensen et al., 1973). Within human capital, Cunha and Heckman (2007) introduced the CES technology of skill formation, and Cunha et al. (2010) estimated multistage CES production functions in which the substitution of interest is between *investments* (parental time, expenditure) and *existing skill stocks* across developmental stages. Attanasio et al. (2020) extended this approach using a randomised controlled trial (RCT) in Colombia. These papers estimate how investments and skills combine over time, but they do not address the question posed here: how do *cognitive and noncognitive skills themselves* combine as co-inputs in the production of a specific educational outcome?

1.1.2 Noncognitive skills and achievement

A large body of work establishes that both cognitive and noncognitive skills predict educational attainment, labour market success, and broader life outcomes (Almlund et al., 2011; Heckman et al., 2006; Kautz et al., 2014). Noncognitive skills - conscientiousness, self-regulation, perseverance - rival IQ in predictive power for academic performance (Bowles & Gintis, 2002; Duckworth & Seligman, 2005). A meta-analysis of over 70,000 observations finds that conscientiousness correlates with grades at $r \approx 0.2-0.3$, largely independently of intelligence (Poropat, 2009). Jackson (2018) shows that teacher effects

on noncognitive outcomes predict long-run success above and beyond test scores, and Deming (2017) documents growing labour-market returns to social skills, with the strongest employment and wage growth in jobs requiring both high cognitive and high social skills, which is consistent with complementarity between skill types.

The productivity of each skill type may also depend on the level of the other: dynamic skill-formation models predict that cognitive and noncognitive inputs are cross-productive, with early noncognitive development raising the return to later cognitive investment (Borghans et al., 2008; Cunha & Heckman, 2008). Gil-Hernández (2021) examines whether cognitive and noncognitive skills act as complements or substitutes in educational track placement using German panel data, finding that high-socioeconomic-status (SES) students can partially compensate for low cognitive ability with conscientiousness. However, that study uses interaction terms in a linear framework rather than estimating a formal production function or elasticity of substitution. To the best of my knowledge, no published study estimates the Hicks elasticity of substitution between cognitive and noncognitive skills as co-inputs in a school achievement production function using a flexible functional form. This paper fills that gap.

1.1.3 Gender differences and measurement

Gender differences in noncognitive skills are well documented. Girls show stronger self-discipline even at comparable IQ levels (Duckworth & Seligman, 2005), and boys are more susceptible to behavioural problems, especially in disadvantaged environments (Bertrand & Pan, 2013). Whether these differences translate into different production technologies - different substitution elasticities by gender - is an open question that this paper addresses.

Measurement remains a challenge, as noncognitive constructs span personality traits, behavioural indicators, and economic preferences, with no single consensus measure (Borghans et al., 2008; Humphries & Kosse, 2017). The SDQ (Goodman, 1997) captures school-proximal behaviours with reasonable internal consistency ($\alpha \approx 0.7$ – 0.8 per subscale), while the TIPI (Gosling et al., 2003) covers the Big Five personality dimensions with only two items per trait ($\alpha \approx 0.4$ – 0.6). This paper uses both instruments and treats the contrast between them as informative about how measurement choice shapes production-function inference. A fuller review of the broader noncognitive literature, including malleability across developmental stages and definitional debates, is provided in Appendix 6.1.

2 Data

2.1 Growing Up in Ireland

The analysis uses the Child Cohort ('98) of the Growing Up in Ireland (GUI) survey, a large national longitudinal study that follows individuals from childhood into early adulthood. I draw predictors from Wave 2 (age 13) and Junior Certificate outcomes reported at Wave 3 (age 17–18), with Wave 1 (age 9) variables used for baseline controls and selection diagnostics. Cohort '98 began with 8,568 children, retained 7,525 at Wave 2 and 6,216 at Wave 3, and remains one of the most informative Irish sources for linking early skill profiles to later educational performance.

For estimation, I use a complete-case longitudinal subsample in which individuals have non-missing values for the variables required by each model block. This choice is deliberate: it keeps sample composition fixed within each comparison and avoids reading model differences that are actually sample differences. The trade-off is external validity, since the estimand becomes conditional associations in an analytically consistent subsample rather than population-average effects for the full GUI cohort. Because the linear models include socioeconomic and school controls that have their own missingness, the ordinary least squares (OLS) sample ($N = 3,786$) is smaller than the translog sample ($N = 5,631$), which requires only cognitive, noncognitive, and outcome variables. Section 4.2.2 confirms that this difference does not drive the results.

Event	Date	Age (in years)	Variables of interest
Study-child is born	Nov/97 - Oct/98	0	
Wave 2 data collection	Aug/11 - Mar/12	13	Independent variables: Cognitive variables, SDQ and TIPI scales, controls
Study-child sits the Junior Cert	Jun/13 - Jun/14	15-16	
Wave 3 data collection	Apr/15 - Aug/16	17-18	Dependent variables: Junior Cert scores in Maths and English

Table 1: Timing structure of the analysis in GUI Cohort '98. Predictors are measured at Wave 2 (age 13), while Junior Certificate outcomes (sat at ages 15–16) are reported at Wave 3 (ages 17–18), creating forward temporal ordering between inputs and achievement.

Table 1 summarises this structure. All predictors - cognitive tests, noncognitive measures, and controls - are drawn from Wave 2 (age 13), while Junior Certificate outcomes are reported at Wave 3 (ages 17–18), two to three years later. This timing reduces reverse-causality concerns, although it does not by itself identify causal effects.

Cognitive ability is measured by three Wave 2 tests: the Drumcondra Verbal Reasoning and Numerical Ability assessments, and the British Ability Scales (BAS) Matrices test of

non-verbal reasoning. I combine these into a single principal component (PC), which I use throughout as the cognitive input for both Maths and English outcomes.¹ Using a single cognitive composite to predict achievement across subjects follows the standard approach in the education production-function literature, where multiple test scores serve as indicators of one latent cognitive factor (Attanasio et al., 2020; Cunha et al., 2010). This choice is supported by psychometric evidence that latent correlations between mathematical and verbal ability typically exceed $r = 0.80$ (Kaufman et al., 2012), consistent with a strong general-ability factor in the Cattell-Horn-Carroll (CHC) framework.² Because my research question concerns the complementarity between cognitive and noncognitive inputs rather than domain-specific cognitive trajectories, this single-factor specification is the appropriate one.³

The Junior Certificate (JC) is a national examination taken by most Irish students at ages 15–16, covering mandatory and elective subjects across levels. I map reported grades to the standardised 12-point Overall Performance Scale (OPS), following established practice (Sofroniou et al., 2000), so achievement can be compared consistently across subjects and levels. Because the syllabus is set years in advance, this outcome reflects sustained preparation and therefore provides a natural setting for studying cognitive and behavioural inputs jointly.

2.1.1 Sample flow

Table 2 summarises how the estimation samples are constructed from the original GUI cohort. Survey attrition between waves accounts for the largest reduction; the remaining loss is analyst-imposed through the complete-case requirement described above. The translog production function models use the larger sample ($N = 5,631$ for SDQ specifications; $N = 5,625$ for TIPI, which has six additional missing values). The OLS specifications with full socioeconomic and school controls restrict the sample further ($N = 3,786$ for SDQ; $N = 3,783$ for TIPI). Section 4.2.2 reports a robustness check confirming that this sample difference does not drive the complementarity finding.

¹The three tests capture verbal fluency, vocabulary comprehension, and numerical ability. Details of the individual assessments are provided in Appendix 6.1.

²The CHC framework is the dominant psychometric model of intelligence structure, integrating Cattell’s (1963) fluid–crystallised distinction, Horn’s (1965) expansion to multiple broad abilities, and Carroll’s (1993) three-stratum hierarchy derived from a comprehensive factor-analytic review. The key implication for production-function work is that verbal, numerical, and non-verbal reasoning tests load heavily on a common general factor (g), so combining them into a single composite is not an ad hoc simplification but reflects the empirical structure of cognitive ability.

³An alternative approach, developed by Agostinelli and Wiswall (2024), estimates separate production functions for maths and reading latent skills. This requires enough domain-specific measurements at each wave to identify each factor separately, which the changing test batteries across GUI waves do not straightforwardly support.

Table 2: Sample flow from the original GUI cohort to model-specific estimation samples. Survey attrition drives the first-stage reduction, while later reductions are analyst-imposed complete-case requirements; TIPI specifications are slightly smaller because of additional item missingness.

Stage	<i>N</i>	% of Wave 1
GUI '98 Cohort, Wave 1 (age 9)	8,568	100.0
Retained at Wave 2 (age 13)	7,525	87.8
Retained at Wave 3 (age 17–18)	6,216	72.6
<i>SDQ analytical samples</i>		
Complete case: cognitive + SDQ + JC outcomes	5,631	65.7
+ Complete SES and school controls	3,786	44.2
<i>TIPI analytical samples</i>		
Complete case: cognitive + TIPI + JC outcomes	5,625	65.7
+ Complete SES and school controls	3,783	44.1
<i>Gender subgroups (full SDQ sample)</i>		
Boys	2,667	-
Girls	2,801	-
Gender missing	163	-

2.2 Variables

I estimate separate linear production functions for Maths and English, regressing Junior Certificate OPS scores on cognitive ability, noncognitive measures, their interactions, and a set of controls. The dependent variables, independent variables, and controls are summarised below; Table 3 provides full descriptive statistics.

2.2.1 Dependent variables

The outcomes are Junior Certificate Maths and English scores on the 12-point OPS scale described above.

2.2.2 Independent variables

The key inputs are cognitive ability (the PC composite) and noncognitive skills, measured by two instruments. The Strengths and Difficulties Questionnaire (SDQ) captures behavioural and emotional characteristics through four subscales. The Ten Item Personality Inventory (TIPI) assesses the Big Five personality dimensions. I estimate separate specifications using each instrument, which allows me to test whether production-function parameters are sensitive to the noncognitive measure employed.

Table 3: Descriptive statistics for outcomes, skill inputs, and controls in the main analytical sample. Outcomes are Junior Certificate OPS points, cognition is based on Wave 2 test components, and SDQ variables are recoded so higher values indicate stronger positive traits.

Variable	Mean	Std. Dev.	Min	Max	N
Dependent variables					
Maths points (Junior Cert)	9.60	1.74	2.00	12.00	5631
English points (Junior Cert)	10.15	1.34	5.00	12.00	5631
Independent variables: Cognition					
Drumcondra Verbal Reasoning (% correct answers)	64.89	21.92	0.00	100.00	5631
Drumcondra Numerical Ability (% correct answers)	55.05	22.53	0.00	100.00	5631
Matrices (BAS)	116.68	18.03	10.00	161.00	5631
Cognition (PC, z-score)	0.14	1.33	-4.25	3.32	5631
Cognition (shifted, mean = 100)	100.00	15.00	36.25	136.40	5631
Independent variables: Noncognition (SDQ)					
Emotional resilience	8.29	1.87	0.00	10.00	5631
Good conduct	8.97	1.31	0.00	10.00	5631
Focused behaviour	7.56	2.26	0.00	10.00	5631
Positive peer relationships	8.96	1.41	0.00	10.00	5631
Independent variables: Noncognition (TIPI)					
Agreeable	5.01	1.95	0.50	7.00	5631
Conscientious	4.33	2.07	0.50	7.00	5631
Emotional stability	4.40	1.99	0.50	7.00	5631
Extravert	3.98	1.98	0.50	7.00	5629
Openness	4.73	1.83	0.50	7.00	5627
Controls (SES characteristics)					
Gender (Male = 1)	0.49	0.50	0.00	1.00	5468
Primary caregiver education level	3.97	1.24	1.00	6.00	5631
Secondary caregiver education level	3.86	1.36	1.00	6.00	4440
Income quintile (equivalized)	3.33	1.39	1.00	5.00	5241
Controls (School characteristics, binary)					
DEIS	0.12	0.33	0.00	1.00	5452
Fee-paying	0.10	0.30	0.00	1.00	5452
Mixed-school	0.54	0.50	0.00	1.00	5317

Note: TIPI scale scores on a 1-7 scale in intervals of 0.5, and the original SDQ scales, ranging from 0 to 10, have been inverted (higher scores typically indicate more problems in the original SDQ scale). "Cognition (PC, z-score)" is the first principal component of the three cognitive tests, standardised to mean 0 and SD 1; it is used in the linear OLS specifications. "Cognition (shifted, mean = 100)" is the same PC score rescaled to mean 100 and SD 15 (following the psychometric convention) and is used in the translog and other nonlinear production function specifications, where all inputs must be strictly positive. Education levels are coded from 1 (Primary or less) to 6 (Postgraduate/Higher degree) in the Growing Up in Ireland caregiver questionnaire. The mean values for both primary (3.97) and secondary (3.86) caregivers indicate an average education level between Leaving Certificate and Diploma/Certificate, indicating a higher proportion of educated caregivers in the sample. Income is reported in quintiles, where 1 represents the lowest 20% and 5 the highest 20% of incomes. The mean of 3.33 shows that the sample is slightly skewed towards higher income levels, with families on average being just above the median income quintile. The sample includes 12% DEIS schools (schools in disadvantaged areas), 10% fee-paying schools, and 54% mixed-gender schools. The sample includes a diverse range of school types, characterised by a high proportion of fee-paying schools and a relatively low share of DEIS schools.

2.2.3 Controls

Control variables enter in two blocks. The socioeconomic block includes gender, parental education (primary and secondary caregiver), and equivalised income quintiles (Davis-Kean, 2005; Sirin, 2005). The school block includes mixed-gender status, Delivering Equality of Opportunity in Schools (DEIS) designation, and fee-paying status (OECD, 2012; Pahlke et al., 2014; Smyth et al., 2015). This structure allows the results to separate skill associations from family and school context.

2.3 Noncognitive measurement instruments

2.3.1 Strengths and Difficulties Questionnaire (SDQ)

The SDQ (Goodman, 1997) is a 20-item instrument measuring behavioural and emotional characteristics through four difficulty subscales (each scored 0–10): Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, and Peer-relationship Problems. I invert the scales so that higher values indicate more positive traits - renaming them Emotional Resilience, Good Conduct, Focused Behaviour, and Positive Peer Relationships, respectively - so that positive coefficients have a consistent interpretation throughout. The SDQ was completed by the primary caregiver at Wave 2 (age 13), contemporaneous with the cognitive assessments⁴. The full questionnaire and scoring are reproduced in Appendix 6.3.

The SDQ captures school-proximal behavioural regulation: the subscales directly reflect classroom-relevant traits such as attention, conduct, and emotional control. This makes it a natural candidate for measuring the noncognitive input in an educational production function.

2.3.2 Ten Item Personality Inventory (TIPI)

The TIPI (Gosling et al., 2003) is a brief instrument assessing the Big Five personality dimensions - Openness, Agreeableness, Conscientiousness, Extraversion, and Emotional Stability - on a seven-point scale. In the main specifications, I use Wave 2 primary-caregiver reports (age 13), which are contemporaneous with the SDQ and cognitive measures and precede the outcome by two to three years. Wave 3 reports (including the child's self-report and secondary-caregiver report) are available, correlate well with Wave 2 scores, and are reserved for supplementary robustness checks. The full questionnaire and scoring are reproduced in Appendix 6.2.

⁴In the GUI '98 Cohort, the primary caregiver is self-identified as the person who provides most care and knows most about the child. In the full Wave 1 cohort ($N = 8,568$), 97.5% of primary caregivers are the biological mother (a further 1.2% are other female guardians); the figure is virtually identical in the analytical sample (97.7%). All "primary caregiver" reports in this chapter should therefore be understood as maternal assessments.

The TIPI captures broader personality traits that are less directly tied to classroom behaviour than the SDQ. Comparing results across the two instruments is central to this paper: it tests whether the complementarity finding depends on the specificity of the noncognitive measure.

2.4 Sample construction and attrition

Table 2 traces the sample from the full analytical dataset through to the estimation samples used in different specifications. The nonlinear production function models (translog, Cobb-Douglas, CES) require only non-missing cognitive, noncognitive, and outcome data and retain the full $N = 5,631$. The linear OLS models with full controls additionally require non-missing SES and school variables, reducing the sample to $N \approx 3,786$. This attrition is analyst-imposed (complete-case requirement) rather than survey non-response, and the robustness check in Table 22 confirms that the translog estimates are stable across both sample definitions.

The retained analytical sample is positively selected on prior achievement and family resources: retained children scored higher on Wave 1 reading ($d = +0.35$) and maths ($d = +0.34$) and came from higher-income families ($d = +0.23$). Within the analytical sample, the further restriction to complete controls ($N \approx 3,786$) is weakly patterned on observables (logit pseudo- $R^2 = 0.014$). Full baseline-balance diagnostics, IPW sensitivity, and sample-flow details are reported in Appendix 6.13. The selection pattern is more likely to attenuate than exaggerate complementarity at lower skill levels.

2.5 Limitations

Three categories of limitation shape interpretation of the results. First, identification: the design does not identify causal effects. Although predictors are measured two to three years before the outcome, which reduces concerns about reverse causality, an omitted-variable bias remains possible if unobserved factors are correlated with both skill inputs and Junior Certificate outcomes. I impose a complete-case requirement across waves and variables, which strengthens comparability across specifications but means that the estimand is conditional associations within a selected longitudinal subsample rather than population-average effects for the full cohort. The selection diagnostics (Table 20) and IPW re-estimation (Table 21) in Appendix 6.13 confirm that this restriction does not drive the main findings, but it does limit generalisability.

Second, measurement: both noncognitive instruments have known limitations. The TIPI's brevity (ten items covering five personality dimensions) may limit its ability to capture nuanced traits, which may partly explain the lower precision of TIPI-based elasticity-of-substitution (ES) estimates relative to SDQ-based ones. The SDQ, while more

detailed, relies on a single informant (primary caregiver at Wave 2) and may be subject to reporting bias. Two risks follow: caregivers may rate children with stronger school performance more favourably (differential reporting), and caregiver-specific optimism/pessimism can induce within-family common-source variation across SDQ dimensions. Under classical measurement error, noncognitive coefficients and interaction terms are attenuated toward zero and the ES is biased toward unity.⁵ Because the complementarity finding ($ES < 1$) is the direction *away* from the attenuation bias, the classical measurement error in parent-rated noncognitive skills acts as a one-sided bound: the true degree of complementarity is at least as strong as, and likely stronger than, the estimates reported here. The TIPI-SDQ comparison is also informative against a pure-noise interpretation: the differences are not only in confidence-interval width, but also in point estimates (for example, girls-TIPI-Maths ES above unity), which is consistent with construct differences rather than precision alone. The availability of multi-informant TIPI reports at Wave 3 (primary caregiver, secondary caregiver, and child self-report) provides a partial check: scores correlate well across informants, which supports the construct validity.

Finally, inference: school identifiers are available under restricted access but were not incorporated in this analysis, so school-clustered variance estimation is not implemented in the headline tables. For the translog-derived objects (ES, MRTS, marginal products, output elasticities), I rely on bootstrap standard errors and interpret threshold-based significance conservatively, placing weight on sign, magnitude, and stability across specifications rather than on individual p -values. A design-effect sensitivity exercise (Section 4.2.2) confirms that the main complementarity finding is robust to plausible clustering adjustments.

3 Linear production function

3.1 Model specification

I begin with a linear production function that relates Junior Certificate achievement to cognitive ability, noncognitive measures, and their interactions. The interaction terms are the key objects of interest: they capture whether the marginal return to cognition depends on the level of noncognitive skills, and vice versa. A negative interaction coefficient implies Edgeworth substitutability - the marginal effect of one input declines as the other increases - while a positive coefficient implies Edgeworth complementarity.

The specification is:

⁵Intuitively, measurement error flattens the curvature of the estimated isoquants, making the two inputs appear more substitutable than they truly are. This follows from standard errors-in-variables results applied to the translog interaction and squared terms; see Bound et al. (2001) for a general treatment.

$$\begin{aligned} \text{PointsJC}_{i,l} = & \beta_0 + \beta_C \cdot \text{Cognition}_i + \sum_{j=1}^J \beta_{Nj} \cdot \text{NonCog}_{i,k,j} \\ & + \sum_{j=1}^J \gamma_j \cdot (\text{Cognition}_i \cdot \text{NonCog}_{i,k,j}) + \boldsymbol{\delta}' \cdot \mathbf{X}_i + \varepsilon_{i,l} \end{aligned} \quad (1)$$

Where i = individual observation, l = Subject (Maths, English), k = Primary caregiver (PCG), j indexes noncognitive skill dimensions (e.g., SDQ and TIPI traits), β_C = coefficient for cognitive ability, β_{Nj} = coefficient for the j -th noncognitive measure, γ_j = coefficient for the interaction between cognition and the j -th noncognitive measure, $\boldsymbol{\delta}'$ = vector of coefficients for control variables.

3.2 Estimation

The variable definitions follow directly from Section 2: Cognition_i is the PC composite, $\text{NonCog}_{i,k,j}$ indexes the four SDQ subscales (or five TIPI traits) reported by the primary caregiver, and \mathbf{X}_i includes the socioeconomic and school controls. All skill inputs are standardised to mean zero and unit standard deviation, so coefficients represent the change in OPS points associated with a one-standard-deviation increase in the input.

I estimate three models for each subject: Model 1 includes only cognitive and noncognitive inputs; Model 2 adds socioeconomic and school controls; and Model 3 adds interaction terms between cognition and each noncognitive measure. This stepwise structure shows how the skill associations change once family background and school context are accounted for, and whether the interaction terms reveal meaningful cross-effects. Separate specifications using SDQ and TIPI allow me to assess whether the results depend on the noncognitive instrument.

The linear specification has well-known limitations: it imposes constant returns to scale and additive main effects, and it does not account for measurement error in skill assessments. These motivate the translog extension in Section 4, which relaxes the functional-form restrictions and allows elasticity of substitution to vary with input levels.

Given these constraints, robustness is treated as a stability exercise rather than a claim of identification. The key criterion is whether signs, relative magnitudes, and core interaction patterns remain consistent across reasonable alternative specifications and sample definitions. Where significance changes at the margin, I prioritise the direction and economic size of effects over p-value thresholds alone.

3.3 Estimates

Table 4: OLS associations between cognition, TIPI personality traits, and Junior Certificate achievement in Maths and English. Dependent variables are OPS points (Maths 2–12; English 5–12); predictors are standardised; Model (1) includes cognition and TIPI only, Model (2) adds socioeconomic and school controls, and Model (3) adds cognition-trait interactions. Standard errors are reported in parentheses and estimates are descriptive rather than causal.

Model:	Maths			English		
	(1)	(2)	(3)	(1)	(2)	(3)
<i>Variables</i>						
Constant	9.1703*** (0.0480)	9.4124*** (0.0633)	9.3992*** (0.0638)	9.9510*** (0.0423)	10.2783*** (0.0533)	10.2705*** (0.0538)
Cognition	0.8280*** (0.0129)	0.7130*** (0.0167)	0.8777*** (0.0428)	0.4993*** (0.0122)	0.4437*** (0.0148)	0.5300*** (0.0399)
Agreeableness	0.0090 (0.0182)	0.0286 (0.0216)	0.0306 (0.0220)	0.0272* (0.0160)	0.0348* (0.0182)	0.0365** (0.0186)
Conscientiousness	0.1609*** (0.0181)	0.1349*** (0.0208)	0.1397*** (0.0215)	0.0900*** (0.0155)	0.0751*** (0.0176)	0.0791*** (0.0182)
Emotional stability	0.0661*** (0.0182)	0.0474** (0.0211)	0.0567*** (0.0219)	0.0135 (0.0160)	0.0046 (0.0178)	0.0087 (0.0185)
Extraversion	-0.0177 (0.0175)	-0.0206 (0.0201)	-0.0185 (0.0208)	0.0164 (0.0151)	0.0114 (0.0168)	0.0135 (0.0174)
Openness	-0.0752*** (0.0175)	-0.0225 (0.0203)	-0.0238 (0.0209)	0.0063 (0.0158)	0.0113 (0.0176)	0.0141 (0.0182)
Male		-0.1142*** (0.0397)	-0.1160*** (0.0397)		-0.4452*** (0.0344)	-0.4440*** (0.0344)
Cognition × Agree.			-0.0168 (0.0166)			-0.0118 (0.0151)
Cognition × Consc.			-0.0168 (0.0159)			-0.0176 (0.0146)
Cognition × Emot.			-0.0669*** (0.0159)			-0.0264* (0.0148)
Cognition × Extra.			0.0011 (0.0155)			-0.0063 (0.0139)
Cognition × Open.			0.0074 (0.0162)			-0.0240* (0.0142)
<i>Controls</i>						
SES	No	Yes	Yes	No	Yes	Yes

Continued on next page

Table 4 continued

Model:	Maths			English		
	(1)	(2)	(3)	(1)	(2)	(3)
School	No	Yes	Yes	No	Yes	Yes
<i>Fit statistics</i>						
Observations	5,625	3,783	3,783	5,625	3,783	3,783
R ²	0.45483	0.48360	0.48732	0.27840	0.33049	0.33361
Adjusted R ²	0.45425	0.48182	0.48487	0.27763	0.32818	0.33042

*Heteroskedasticity-robust (HC1) standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Agree.: Agreeableness, Consc.: Conscientiousness, Emot.: Emotional stability, Extra.: Extraversion, Open.: Openness. Note: Regressors are standardised (mean 0, SD 1). Coefficients are reported in outcome units (points) per 1 SD increase in the predictor.*

3.4 Linear results

Tables 4 and 5 report the linear results. Three patterns organise the interpretation.

First, cognitive ability dominates: a one-SD increase in cognition is associated with 0.72–0.84 points in Maths and 0.45–0.50 points in English, making it the strongest predictor in every specification. Since one OPS point corresponds to one grade band, this implies roughly two-thirds of a grade-band shift in Maths for a one-SD change in cognition, with a smaller but still substantive effect in English.

Second, Focused Behaviour is the strongest noncognitive predictor. Under the pre-committed hierarchy (SDQ Focused Behaviour primary, TIPI Conscientiousness confirmatory), a one-SD increase in Focused Behaviour is associated with 0.23 points in Maths and 0.18 points in English, whereas Conscientiousness shows smaller associations (0.14 and 0.08). The noncognitive margin is smaller than the cognitive margin but still economically relevant, and its explanatory contribution is larger in English than in Maths.

Third, the interaction terms are predominantly negative. The Cognition \times Focused Behaviour interaction is negative and significant in both subjects (-0.057 in Maths, -0.074 in English), indicating Edgeworth substitutability in marginal effects.⁶ For TIPI, the corresponding interactions are weaker and less precise. These cross-partial patterns motivate the translog extension, which tests whether low substitutability also holds along isoquants in the Hicks sense.

⁶Edgeworth substitutability means that the marginal productivity of one input falls as the level of the other rises - i.e., the cross-partial derivative $\partial^2 Y / \partial C \partial N < 0$. This is distinct from Hicks substitutability, which asks how easily one input can *replace* the other along an isoquant (a curve of constant output). A production function can exhibit Edgeworth substitutability (negative interaction) and Hicks complementarity ($ES < 1$) simultaneously; Section 4.1 develops the distinction formally.

Table 5: OLS associations between cognition, SDQ behavioural characteristics, and Junior Certificate achievement in Maths and English. Dependent variables are OPS points (Maths 2–12; English 5–12); predictors are standardised; Model (1) includes cognition and SDQ only, Model (2) adds socioeconomic and school controls, and Model (3) adds cognition-trait interactions. Standard errors are reported in parentheses and estimates are descriptive rather than causal.

Model:	Maths			English		
	(1)	(2)	(3)	(1)	(2)	(3)
<i>Variables</i>						
Constant	8.5629*** (0.1085)	8.7291*** (0.1314)	8.8320*** (0.1294)	9.4481*** (0.0992)	9.7286*** (0.1150)	9.8366*** (0.1123)
Cognition	0.7763*** (0.0135)	0.6675*** (0.0172)	0.9811*** (0.0846)	0.4595*** (0.0124)	0.4135*** (0.0149)	0.7233*** (0.0807)
Emotional Resilience	0.0507** (0.0198)	0.0544** (0.0233)	0.0444* (0.0228)	-0.0258 (0.0179)	0.0092 (0.0206)	-0.0007 (0.0203)
Good Conduct	0.0832*** (0.0207)	0.0693*** (0.0250)	0.0714*** (0.0246)	0.0060 (0.0191)	-0.0201 (0.0221)	-0.0161 (0.0215)
Focused Behaviour	0.2447*** (0.0210)	0.2283*** (0.0249)	0.2175*** (0.0247)	0.2440*** (0.0187)	0.1906*** (0.0218)	0.1782*** (0.0213)
Positive Peer Relationships	0.0197 (0.0181)	0.0002 (0.0211)	0.0046 (0.0224)	0.0704*** (0.0170)	0.0502** (0.0199)	0.0565*** (0.0208)
Male		-0.0707* (0.0398)	-0.0702* (0.0397)		-0.4019*** (0.0342)	-0.3996*** (0.0339)
Cognition × E.R.			-0.0283* (0.0163)			-0.0148 (0.0154)
Cognition × G.C.			0.0122 (0.0181)			0.0247 (0.0180)
Cognition × F.B.			-0.0567*** (0.0174)			-0.0737*** (0.0171)
Cognition × P.P.R.			-0.0158 (0.0144)			-0.0261* (0.0147)
<i>Controls</i>						
SES	No	Yes	Yes	No	Yes	Yes
School	No	Yes	Yes	No	Yes	Yes
<i>Fit statistics</i>						
Observations	5,631	3,786	3,786	5,631	3,786	3,786
R ²	0.47270	0.49904	0.50214	0.30700	0.34584	0.35277
Adjusted R ²	0.47223	0.49744	0.50002	0.30639	0.34376	0.35003

*Heteroskedasticity-robust (HC1) standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. E.R.: Emotional Resilience, G.C.: Good Conduct, F.B.: Focused Behaviour, P.P.R.: Positive Peer Relationships*

Note: Regressors are standardised (mean 0, SD 1). Coefficients are reported in outcome units (points) per 1 SD increase in the predictor.

4 Translog production function

The linear specification establishes that cognitive and noncognitive skills are both associated with achievement and that their interaction terms are predominantly negative (Edgeworth substitutability). But a linear model cannot answer the deeper economic question: how easily can one skill type *replace* the other at constant achievement? The Allen-Uzawa elasticity of substitution (ES)⁷ directly addresses this, and computing it requires a flexible functional form.

I use the translog production function (Christensen et al., 1973), a second-order Taylor approximation in logarithms that nests the Cobb-Douglas as a special case and allows ES to vary with input levels.⁸ In this chapter, the translog is used as a flexible local approximation to the conditional mean function, not as a literal structural technology for bounded exam scores. Junior Certificate OPS grades are discrete and bounded (2–12 points), so the log-log specification should be understood as a convenient parameterisation that disciplines the curvature of the input–output relationship rather than as a claim that educational achievement is generated by a smooth, unbounded production process.⁹ I model achievement as a function of two inputs: cognition (C) and a single noncognitive measure (N). To reduce researcher degrees of freedom, I pre-commit to two specifications: SDQ Focused Behaviour as the primary noncognitive input (chosen because sustained attention and task persistence are the dimensions most directly relevant to examination performance; (Duncan et al., 2007)), and TIPI Conscientiousness as confirmatory (a broader personality construct that overlaps with but is not identical to school-relevant behavioural regulation).

4.1 Defining complementarity

Before presenting the specification and results, it is important to clarify how complementarity is used in this paper, since the linear and translog models capture different facets of the relationship between skills.

A concrete example helps fix intuition. Consider a student with high cognitive ability.

⁷The Allen-Uzawa ES measures how easily one input can replace another along an isoquant (i.e., holding output constant). $ES = 1$ corresponds to Cobb-Douglas (unit substitutability); $ES < 1$ means the inputs are hard to trade off (Hicks complements); $ES > 1$ means they are easily substituted. With two inputs, the Allen–Uzawa ES coincides with the Morishima and Hicks elasticities, so the choice among definitions is immaterial here.

⁸Appendix results for the more restrictive Cobb-Douglas and CES specifications confirm the qualitative pattern: cognition is the dominant input, noncognitive contributions are smaller but meaningful, and substitution elasticities are below one. Direct CES estimation in this dataset was numerically less stable in several configurations (including convergence failures and boundary solutions), so the translog is preferred because it identifies substitutability from local curvature while relaxing constant-ES restrictions.

⁹Sections 4.2.2 and the robustness tables confirm that the key ES result survives under z-scored outcomes, top-decile trimming, and alternative evaluation points, none of which depend on the log transformation.

Adding more focused behaviour still helps, but the additional gain is smaller than it would be for a lower-ability student—the marginal boost from one input falls as the other rises (Edgeworth substitutability). At the same time, if we ask how much extra behaviour is needed to replace a lost unit of cognition while keeping the student’s grade unchanged, the answer is “a lot”: cognition and behaviour are hard to trade off at constant output (Hicks complementarity, $ES < 1$). Both statements can be true simultaneously, and in this paper both hold in the data.

In the linear models, a negative interaction coefficient γ_j indicates Edgeworth substitutability: the marginal benefit of one input diminishes when the other is already high. In the translog, $ES < 1$ indicates Hicks complementarity: maintaining the same achievement level requires the two inputs to move roughly in tandem, and substituting one for the other is costly. These are different objects and can coexist. The primary concept of interest in this paper is Hicks complementarity ($ES < 1$), which speaks directly to whether cognitive and noncognitive skills should be developed jointly or can be targeted in isolation.

4.2 Specification

The translog function is defined as:

$$Y = AC^\alpha N^\beta \exp \left\{ \frac{1}{2} \gamma_1 [\ln(C)]^2 + \frac{1}{2} \gamma_2 [\ln(N)]^2 + \gamma_{12} \ln(C) \ln(N) \right\} \quad (2)$$

Where:

Y : Total output/Grade function/Academic achievement

A : Total factor productivity or scaling factor

C, N : Inputs

α, β : Exponents determining the output response to each input

$\gamma_1, \gamma_2, \gamma_{12}$: Parameters capturing interactions and quadratic effects

The parameters α and β capture first-order output elasticities; γ_1 and γ_2 capture curvature (diminishing or increasing returns to each input); and γ_{12} captures the cross-input interaction - how the marginal product of one input depends on the level of the other. From these parameters, I compute four production-theory objects at sample means: marginal products (MPs), output elasticities (OEs), the marginal rate of technical substitution (MRTS), and the elasticity of substitution. Full derivations are provided in the Appendix; the key expression for ES is:

$$\sigma = \frac{OE_C + OE_N}{OE_C + OE_N - \gamma_{12} \left(\frac{OE_C}{OE_N} + \frac{OE_N}{OE_C} \right)} \quad (3)$$

where $OE_C = \alpha + \gamma_1 \ln C + \gamma_{12} \ln N$ and $OE_N = \beta + \gamma_2 \ln N + \gamma_{12} \ln C$ are the output elasticities evaluated at sample means. The ES is the primary object of interest: $ES < 1$ indicates Hicks complementarity (low substitutability along isoquants), $ES = 1$ corresponds to the Cobb-Douglas case, and $ES > 1$ indicates that the inputs are easy to trade off against each other at constant output.

4.2.1 Results

Table 6: Translog production-function estimates for Junior Certificate Maths (OPS points, 2–12 scale). SDQ Focused Behaviour is the primary noncognitive input; TIPI Conscientiousness is confirmatory. The elasticity of substitution (ES) measures how easily one skill type can replace the other at constant achievement: $ES < 1$ indicates Hicks complementarity (skills must move together); $ES = 1$ is the Cobb-Douglas case; $ES > 1$ indicates easy substitution. Output elasticities (OEs) give the percentage change in achievement for a 1% change in the input. The marginal rate of technical substitution (MRTS) gives the units of noncognitive skill needed to replace one unit of cognition at constant output. All quantities are evaluated at sample means and are descriptive model summaries, not causal effects. $N = 5,631$ (full SDQ sample); $N = 5,625$ (full TIPI sample).

Parameter	Maths			
	Full	Boys	Girls	Δ (B–G)
TIPI Model				
A	9.43*** (0.03)	9.35*** (0.04)	9.51*** (0.04)	-0.16
α (Cognition)	0.83*** (0.02)	0.85*** (0.02)	0.82*** (0.02)	0.03
β (Consc.)	0.04*** (0.01)	0.04*** (0.01)	0.05*** (0.01)	-0.01
γ_1	0.10 (0.11)	0.31* (0.18)	-0.15 (0.16)	0.46
γ_2	0.05*** (0.01)	0.03* (0.02)	0.07*** (0.02)	-0.04
γ_{12}	-0.04 (0.03)	-0.09** (0.04)	0.004 (0.04)	-0.09
MP (Cognition)	0.078	0.079	0.079	0.000
MP (Consc.)	0.097	0.081	0.107	-0.026
OE (Cognition)	0.824	0.848	0.821	0.027
OE (Consc.)	0.049	0.040	0.056	-0.016

Continued on next page

Table 6 continued

Parameter	Maths			
	Full	Boys	Girls	Δ (B–G)
ES	0.569	0.312	1.081	–0.769
MRTS	0.804	0.977	0.736	0.241
SDQ Model				
A	9.45*** (0.02)	9.35*** (0.03)	9.54*** (0.03)	–0.19
α (Cognition)	0.79*** (0.02)	0.81*** (0.02)	0.78*** (0.02)	0.03
β (Foc. Behav.)	0.10*** (0.01)	0.09*** (0.01)	0.12*** (0.01)	–0.03
γ_1	0.20* (0.11)	0.37** (0.18)	0.04 (0.16)	0.33
γ_2	0.06*** (0.01)	0.05*** (0.01)	0.08*** (0.02)	–0.03
γ_{12}	–0.13*** (0.03)	–0.11*** (0.04)	–0.15*** (0.05)	0.04
MP (Cognition)	0.074	0.075	0.074	0.001
MP (Foc. Behav.)	0.128	0.109	0.146	–0.037
OE (Cognition)	0.777	0.803	0.770	0.033
OE (Foc. Behav.)	0.109	0.088	0.129	–0.041
ES	0.482	0.463	0.497	–0.034
MRTS	0.576	0.694	0.508	0.186
Observations	5,631	2,667	2,801	

SEs in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Δ is descriptive.

Table 7: Translog production-function estimates for Junior Certificate English (OPS points, 5–12 scale). SDQ Focused Behaviour is the primary noncognitive input; TIPI Conscientiousness is confirmatory. $ES < 1$ indicates Hicks complementarity (skills must move together for efficient production); $ES = 1$ is the Cobb-Douglas case. Output elasticities (OEs) give the percentage change in achievement for a 1% change in the input. MRTS gives the units of noncognitive skill needed to replace one unit of cognition at constant output. All quantities are evaluated at sample means and are descriptive, not causal. $N = 5,631$ (SDQ); $N = 5,625$ (TIPI).

Parameter	English			
	Full	Boys	Girls	Δ (B–G)
TIPI Model				
A	10.09*** (0.02)	9.84*** (0.04)	10.36*** (0.03)	-0.52
α (Cognition)	0.45*** (0.01)	0.51*** (0.02)	0.44*** (0.02)	0.07
β (Consc.)	0.02*** (0.00)	0.02*** (0.01)	0.02*** (0.01)	0.00
γ_1	-0.25*** (0.09)	-0.25 (0.15)	-0.32*** (0.12)	0.07
γ_2	0.03*** (0.01)	0.03* (0.02)	0.02 (0.01)	0.01
γ_{12}	-0.03 (0.02)	-0.04 (0.03)	-0.01 (0.03)	-0.03
MP (Cognition)	0.046	0.049	0.047	0.002
MP (Consc.)	0.059	0.048	0.039	0.009
OE (Cognition)	0.450	0.501	0.446	0.055
OE (Consc.)	0.028	0.022	0.019	0.003
ES	0.514	0.383	0.698	-0.315
MRTS	0.775	1.027	1.207	-0.180
SDQ Model				
A	10.08*** (0.02)	9.83*** (0.03)	10.34*** (0.03)	-0.51
α (Cognition)	0.41*** (0.01)	0.47*** (0.02)	0.41*** (0.02)	0.06
β (Foc. Behav.)	0.09*** (0.01)	0.07*** (0.01)	0.08*** (0.01)	-0.01
γ_1	-0.17* (0.01)	-0.14 (0.01)	-0.17 (0.01)	0.03

Continued on next page

Table 7 continued

Parameter	English			
	Full	Boys	Girls	Δ (B–G)
	(0.09)	(0.15)	(0.12)	
γ_2	0.06***	0.05***	0.06***	-0.01
	(0.01)	(0.01)	(0.01)	
γ_{12}	-0.12***	-0.12***	-0.15***	0.03
	(0.02)	(0.03)	(0.04)	
MP (Cognition)	0.041	0.045	0.043	0.002
MP (Foc. Behav.)	0.116	0.094	0.100	-0.006
OE (Cognition)	0.406	0.456	0.409	0.047
OE (Foc. Behav.)	0.092	0.073	0.081	-0.008
ES	0.464	0.411	0.381	0.030
MRTS	0.356	0.476	0.430	0.046
Observations	5,631	2,667	2,801	

SEs in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Δ is descriptive.

Table 8: Elasticity of substitution (ES) from translog models with bootstrap inference. ES below 1 is consistent with complementarity; ES above 1 indicates greater substitutability. Confidence intervals are percentile bootstrap (1,000 replications). $p(\text{ES} = 1)$ is a two-sided bootstrap test of the Cobb-Douglas benchmark. Values are model-based descriptive summaries for the estimation sample at noncognitive shift +0.5.

Sample	Outcome	Noncognitive measure	N	ES	ES [95% CI]	$p(\text{ES} = 1)$
Full sample	Maths	SDQ (Focused Behaviour)	5,631	0.482	[0.378, 0.673]	0.000
Full sample	Maths	TIPI (Conscientiousness)	5,631	0.569	[0.354, 1.701]	0.240
Full sample	English	SDQ (Focused Behaviour)	5,631	0.464	[0.349, 0.690]	0.000
Full sample	English	TIPI (Conscientiousness)	5,631	0.514	[0.226, 1.914]	0.180
Boys	Maths	SDQ (Focused Behaviour)	2,667	0.463	[0.284, 0.785]	0.010
Boys	Maths	TIPI (Conscientiousness)	2,667	0.312	[0.181, 0.990]	0.050
Boys	English	SDQ (Focused Behaviour)	2,667	0.411	[0.279, 0.627]	0.000
Boys	English	TIPI (Conscientiousness)	2,667	0.383	[-3.226, 2.011]	0.190
Girls	Maths	SDQ (Focused Behaviour)	2,801	0.497	[0.347, 0.766]	0.010
Girls	Maths	TIPI (Conscientiousness)	2,801	1.081	[-6.403, 7.943]	0.830
Girls	English	SDQ (Focused Behaviour)	2,801	0.381	[0.230, 0.701]	0.010
Girls	English	TIPI (Conscientiousness)	2,801	0.698	[-4.926, 3.554]	0.300

The cognitive output elasticities are large and significant in both subjects ($OE_C \approx 0.78$ – 0.82 for Maths, 0.41 – 0.45 for English), confirming that cognitive ability is the dominant input. The noncognitive elasticities are smaller but consistently positive, with the SDQ measure showing stronger associations ($OE_N = 0.109$ for Maths, 0.092 for English) than

the TIPI (0.049 and 0.028).

The interaction parameter γ_{12} is negative and significant in the SDQ models for both subjects (-0.131 for Maths, -0.124 for English), indicating Edgeworth substitutability at the margin: the marginal product of each input decreases as the other rises. However, the economically more informative object is the elasticity of substitution, which is consistently below one in the SDQ specifications (0.482 for Maths, 0.464 for English; Table 8). This confirms Hicks complementarity: maintaining a given achievement level becomes progressively costly when one input must compensate for the other. The coexistence of negative γ_{12} (Edgeworth) and $ES < 1$ (Hicks) is explained in Section 4.1.

Gender differences in the core parameters are present but moderate: boys show slightly larger cognitive elasticities, girls slightly larger noncognitive elasticities, and ES estimates are comparable across gender within the SDQ block. The one prominent exception is the TIPI girls-Maths point estimate ($ES = 1.081$), but the noncognitive output elasticity in that cell is close to zero, leaving the ES ratio effectively undetermined: bootstrap intervals span $[-6.4, 7.9]$ and do not distinguish complementarity from substitutability. Classical measurement error in parent-rated noncognitive skills would bias the ES *toward* unity, so the finding of $ES < 1$ is conservative (see Section 2.5).

Figure 1 presents the complementarity pattern from the linear interaction models. Each panel shows predicted achievement as a function of cognitive ability, evaluated at low (-1 SD), median, and high ($+1$ SD) noncognitive skill (Focused Behaviour). The diverging slopes confirm that the marginal return to cognition is higher when noncognitive skills are low, and lower when noncognitive skills are already high-consistent with the negative interaction terms reported in the OLS tables and the $ES < 1$ finding from the translog. In practical terms, a cognitively able child who struggles with attention and behavioural regulation is predicted to gain less from that cognitive ability than an equally able child with stronger behavioural skills; high cognition alone does not fully compensate for low noncognitive development.

Figure 2 plots iso-achievement curves from the translog specification in (cognition, noncognition) space. The convex shape of the isoquants confirms $ES < 1$: maintaining a given level of achievement requires increasingly large amounts of one input to compensate for small reductions in the other. The curvature is more pronounced in English than in Maths, matching the lower ES point estimates for English.

To translate the translog estimates into policy-relevant units: at the sample means, a 10% increase in cognition is associated with a 7.8% increase in Maths points (≈ 0.75 OPS points, or three-quarters of a grade band) and a 4.1% increase in English points (≈ 0.41 points). A 10% increase in Focused Behaviour is associated with a 1.1% increase in Maths and 0.9% in English-smaller, but operating through a different and complementary channel. The MRTS estimates provide an intuitive summary of complementarity: compensating for

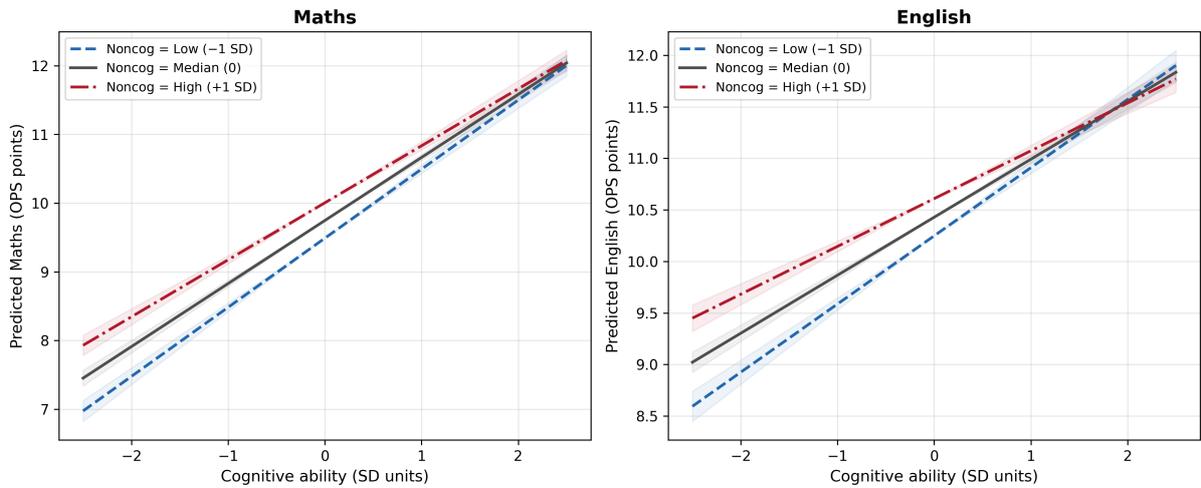


Figure 1: Predicted achievement from the linear interaction model (SDQ Focused Behaviour), holding controls at their means. Each line evaluates the marginal effect of cognitive ability at a different level of noncognitive skill (-1 SD, median, $+1$ SD). The diverging slopes show that the return to cognition diminishes as noncognitive skill increases, consistent with complementarity (negative interaction). Shaded bands are 95% confidence intervals on the marginal effect. $N \approx 3,786$; HC1 robust SEs.

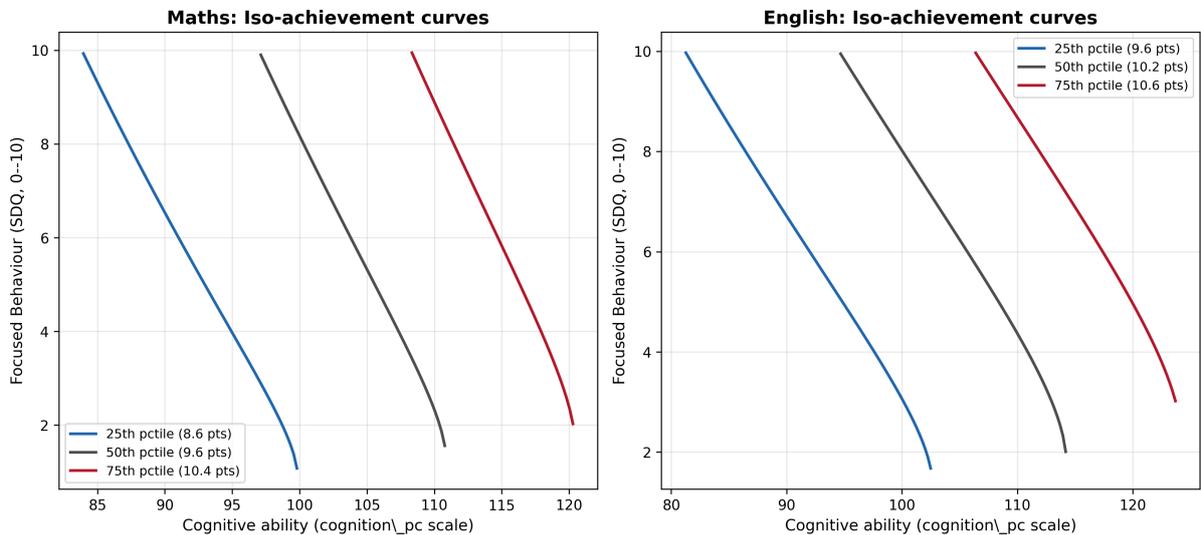


Figure 2: Translog iso-achievement curves for Maths and English (SDQ Focused Behaviour, shift = $+0.5$). Each curve traces combinations of cognitive ability and noncognitive skill that yield the same predicted achievement level (25th, 50th, and 75th percentiles of predicted log points). Convexity confirms $ES < 1$ (complementarity): substituting one input for the other becomes progressively costly.

a one-unit reduction in cognition requires approximately 7 units of additional noncognitive input in Maths and 4 units in English. These magnitudes emphasize why developing both skill types jointly is more efficient than relying on one input to offset deficits in the other.

4.2.2 Inference and sensitivity

For the linear OLS models, I report heteroskedasticity-robust (HC1) standard errors throughout. For the nonlinear production-function estimates, inference relies on a non-parametric bootstrap with 1,000 replications (95% percentile intervals), and the elasticity of substitution is tested against $ES = 1$ with a two-sided bootstrap p -value.

To keep the main argument focused, the core robustness checks are prioritised in the main text and summarised in Table 13: (i) Wave 1 baseline controls, (ii) sample-restriction and IPW sensitivity, (iii) clustering/design-effect sensitivity, and (iv) scale/functional-form sensitivity (including z -score analogues and top-decile trimming). Secondary diagnostics are reported in detailed tables and appendices.

As the preferred baseline-control linear specification¹⁰ I re-estimate the interaction models with Wave 1 (age 9) baseline controls included. These include standardised reading and maths test scores from Wave 1 and four parent-reported SDQ subscales at age 9. By including these measures, I aim to absorb stable child-level traits and early-environment effects that might otherwise confound the association between Wave 2 inputs and Junior Certificate outcomes. Wave 1 baselines are individually significant and increase R^2 by 3–6 percentage points, confirming that age-9 skills carry forward. More importantly, all 18 cognitive \times noncognitive interaction terms preserve their signs after the addition of Wave 1 controls, with a median absolute attenuation of 22%. The key interaction-SDQ Focused Behaviour \times Cognition-moves from -0.056 to -0.052 in Maths ($p < 0.01$ in both) and from -0.073 to -0.063 in English ($p < 0.001$ in both). These results confirm that the estimated interaction patterns are not simply proxying for time-invariant ability or temperament (Table 9).

The translog ES estimates are stable when re-estimated on the restricted complete-controls sample ($N = 3,786$), with point estimates changing by at most 0.04 (Appendix 6.14, Table 22).

The complementarity conclusion does not depend on the log transformation: a second-order polynomial with z -scored outcomes produces the same sign pattern across all 12 specifications (Appendix 6.14, Table 23). Trimming the upper tail of the outcome distribution likewise leaves the primary SDQ ES estimates unchanged (Table 24). Ad-

¹⁰I use “baseline control” rather than “value-added” to describe the inclusion of Wave 1 skill measures, following the taxonomy of Todd and Wolpin (2003): a proper value-added specification requires lagged *same*, outcome scores, whereas the present design conditions on earlier-wave skill inputs that are related to but not identical to Junior Certificate grades. The distinction matters for the interpretation of omitted-variable bias.

Table 9: OLS interaction coefficients with and without Wave 1 baseline controls (preferred baseline-control specification).

Spec	Interaction	Model 4 (+W1, preferred)		Model 3		Change (%)
		Coef	SE	Coef	SE	
<i>Panel A: SDQ (Focused Behaviour) - Primary</i>						
Maths	Emot. Res.×Cog	-0.017	(0.016)	-0.028 ⁺	(0.016)	+38
Maths	Good Cond.×Cog	0.009	(0.018)	0.012	(0.018)	-26
Maths	Foc. Behav.×Cog	-0.052 ^{**}	(0.017)	-0.056 ^{**}	(0.017)	+7
Maths	Pos. Peers×Cog	-0.010	(0.015)	-0.016	(0.014)	+38
English	Emot. Res.×Cog	-0.007	(0.015)	-0.015	(0.015)	+51
English	Good Cond.×Cog	0.020	(0.017)	0.025	(0.018)	-19
English	Foc. Behav.×Cog	-0.063 ^{***}	(0.016)	-0.073 ^{***}	(0.017)	+14
English	Pos. Peers×Cog	-0.020	(0.014)	-0.026 ⁺	(0.015)	+23
<i>Panel B: TIPI (Conscientiousness) - Confirmatory</i>						
Maths	Agreeable×Cog	-0.012	(0.016)	-0.017	(0.017)	+28
Maths	Conscient.×Cog	-0.014	(0.015)	-0.017	(0.016)	+13
Maths	Emo. Stab.×Cog	-0.061 ^{***}	(0.015)	-0.067 ^{***}	(0.016)	+8
Maths	Extravert×Cog	0.004	(0.015)	0.001	(0.015)	-
Maths	Openness×Cog	0.008	(0.016)	0.007	(0.016)	+14
English	Agreeable×Cog	-0.009	(0.015)	-0.012	(0.015)	+20
English	Conscient.×Cog	-0.015	(0.014)	-0.018	(0.015)	+15
English	Emo. Stab.×Cog	-0.020	(0.014)	-0.026 ⁺	(0.015)	+23
English	Extravert×Cog	-0.001	(0.013)	-0.006	(0.014)	+80
English	Openness×Cog	-0.023 ⁺	(0.014)	-0.024 ⁺	(0.014)	+3

Note: Model 4 (preferred) includes Wave 2 cognitive ability (PC), noncognitive dimensions, SES and school controls, cognition×noncognition interaction terms, and standardised Wave 1 baselines (reading score, maths score, and four parent-reported SDQ subscales at age 9). Model 3 omits Wave 1 baselines. All 18 interaction terms preserve their signs. HC1 robust standard errors. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

ditional diagnostic checks such as translog regularity conditions, ES stability across evaluation points and shift values, Oster (2019) bounds for omitted variable bias, and design-effect-inflated inference under plausible school-clustering parameters, are reported in Appendix 6.14.

The analysis estimates multiple specifications (instrument, subject, and subgroup variants), so I pre-commit to a hierarchy: SDQ Focused Behaviour is primary, TIPI Conscientiousness is confirmatory, and broader grids are exploratory. The main conclusion is therefore anchored to the primary SDQ specification and then stress-tested for stability. As a transparency check, I report Benjamini–Hochberg false discovery rate (FDR) q -values for the 18 cognition \times noncognition interaction tests in the baseline-control specifications (Table 10). Of these 18 tests, only three survive the 10% FDR threshold - and two of the three are the primary SDQ Focused Behaviour \times Cognition interactions that anchor the paper’s complementarity finding (Maths $q = 0.012$; English $q = 0.001$). The third survivor is TIPI Emotional Stability \times Cognition in Maths ($q = 0.001$). The signal is concentrated in the theoretically central terms - school-proximal behavioural regulation interacting with cognitive ability - rather than spread diffusely across the full interaction grid.

Table 10: Interaction-grid p -values with Benjamini-Hochberg false-discovery-rate (FDR) adjustment. Because the analysis tests 18 separate cognition \times noncognitive interactions (4 SDQ subscales + 5 TIPI dimensions \times 2 subjects), some may appear significant by chance alone. The q_{BH} column adjusts for this multiple testing: it controls the expected share of false positives among all rejected hypotheses. Baseline-control models (Wave 1 baselines included).

Outcome	Spec	Interaction	Coef	SE	p	q_{BH}
Maths	SDQ	SDQ Emot. Res. \times Cog	-0.0176	0.0165	0.2845	0.5704
Maths	SDQ	SDQ Good Cond. \times Cog	0.0090	0.0180	0.6162	0.7005
Maths	SDQ	SDQ Foc. Behav. \times Cog	-0.0528	0.0171	0.0020	0.0121 [†]
Maths	SDQ	SDQ Pos. Peers \times Cog	-0.0098	0.0147	0.5036	0.7005
Maths	TIPI	Agreeable W2 PCG \times Cog	-0.0121	0.0163	0.4589	0.7005
Maths	TIPI	Conscientious W2 PCG \times Cog	-0.0146	0.0155	0.3486	0.6275
Maths	TIPI	Emo Stability W2 PCG \times Cog	-0.0619	0.0156	0.0001	0.0010 [†]
Maths	TIPI	Extravert W2 PCG \times Cog	0.0037	0.0151	0.8073	0.8548
Maths	TIPI	Openness W2 PCG \times Cog	0.0085	0.0156	0.5885	0.7005
English	SDQ	SDQ Emot. Res. \times Cog	-0.0074	0.0150	0.6226	0.7005
English	SDQ	SDQ Good Cond. \times Cog	0.0200	0.0175	0.2528	0.5704
English	SDQ	SDQ Foc. Behav. \times Cog	-0.0633	0.0164	0.0001	0.0010 [†]
English	SDQ	SDQ Pos. Peers \times Cog	-0.0202	0.0143	0.1601	0.4804
English	TIPI	Agreeable W2 PCG \times Cog	-0.0094	0.0147	0.5215	0.7005
English	TIPI	Conscientious W2 PCG \times Cog	-0.0151	0.0141	0.2852	0.5704
English	TIPI	Emo Stability W2 PCG \times Cog	-0.0203	0.0144	0.1588	0.4804
English	TIPI	Extravert W2 PCG \times Cog	-0.0013	0.0132	0.9214	0.9214
English	TIPI	Openness W2 PCG \times Cog	-0.0233	0.0136	0.0870	0.3914

Note: Rows report cognition \times noncognition interactions from baseline-control OLS models (Wave 1 baselines included). q -values are Benjamini-Hochberg adjusted across all 18 interaction tests. [†] indicates $q_{BH} < 0.10$.

To complement the split-sample presentations, I also estimate pooled baseline-control

models with explicit gender interaction structure and joint Wald tests (Table 11). The null that all gender differences in cognition \times noncognition interactions are zero is not rejected in the four pooled specifications (p-values 0.29–0.65), which supports treating subgroup contrasts as suggestive heterogeneity patterns rather than sharp structural breaks. A power calculation contextualises this null: with $N \approx 1,850$ per gender subgroup and a pooled interaction SE of approximately 0.017 for the focal SDQ term, the minimum detectable effect (MDE) for a gender difference in the interaction coefficient is approximately 0.10 at 80% power ($\alpha = 0.05$).¹¹ Since the main interaction coefficient itself is only -0.052 in Maths, the test is powered to detect gender differences roughly twice the size of the main effect—a large threshold that makes the non-rejection unsurprising and consistent with modest or absent gender heterogeneity in the production technology.

Table 11: Pooled gender-difference joint tests in baseline-control interaction models.

Outcome	Spec	N	p(Joint gender \times interaction = 0)	p(Joint all gender-diff terms = 0)
Maths	SDQ	3711	0.3525	0.3087
Maths	TIPI	3708	0.4430	0.7152
English	SDQ	3711	0.6531	0.6732
English	TIPI	3708	0.2935	0.6727

Note: HC1-robust Wald tests from pooled models with Wave 1 baseline controls. The first p-value tests whether all gender differences in cognition \times noncognition interactions are zero; the second tests all gender-difference terms jointly.

Finally, I stress-test the cognition construct by rebuilding PC under leave-one-subtest-out variants and re-estimating the primary SDQ baseline-control interaction model (Table 12). The key Focused Behaviour \times Cognition interaction remains negative in all variants for both outcomes, with statistical significance preserved at conventional levels (Maths p-values 0.001–0.035; English p-values below 0.001). This supports the conclusion that the main interaction pattern is not an artefact of any single cognitive subtest.

Table 12: Cognition stress test: leave-one-subtest-out PC and key interaction stability.

Outcome	Cognition variant	N	$\hat{\beta}_{FB \times Cog}$	SE	p
Maths	All three subtests (baseline PC)	3711	-0.0522	0.0169	0.0021
English	All three subtests (baseline PC)	3711	-0.0626	0.0162	0.0001
Maths	Drop Drum _V RW2 _p	3711	-0.0518	0.0210	0.0137
English	Drop Drum _V RW2 _p	3711	-0.0619	0.0180	0.0006
Maths	Drop Drum _N AW2 _p	3711	-0.0466	0.0221	0.0350
English	Drop Drum _N AW2 _p	3711	-0.0691	0.0194	0.0004
Maths	Drop BAS _T S _M at _W 2	3711	-0.0610	0.0185	0.0010
English	Drop BAS _T S _M at _W 2	3711	-0.0675	0.0181	0.0002

Note: Primary SDQ baseline-control specification. Each row rebuilds cognition from a different test subset and reports the Focused Behaviour \times Cognition interaction coefficient.

As a further robustness check, I re-estimate the translog production function using Leaving Certificate (LC) points rather than Junior Certificate grades as the outcome.

¹¹Using the approximation $MDE \approx 2.8 \times SE_{diff}$, where $SE_{diff} \approx \sqrt{2} \times SE_{pooled} \times \sqrt{N/N_{sub}} \approx 0.034$.

The Leaving Certificate is the terminal upper-secondary examination in Ireland, taken at age 17-18 (two to three years after the JC); results are converted to a standardised points scale used for university admission.¹² This serves two purposes: it confirms that the complementarity finding is not specific to the JC exam, and it tests whether the result holds over a longer developmental window. I estimate two additional specifications: (i) Wave 2 inputs (age 13) predicting LC points (age 17/18), a 4–5 year gap spanning the full secondary-school period; and (ii) Wave 3 inputs (age 17/18) predicting LC points, a near-contemporaneous specification using different cognitive instruments (semantic fluency, mathematical calculations, and vocabulary tests). Table 31 in Appendix 6.15 reports the full results. The key finding is that SDQ-based ES estimates remain below one in both LC specifications for English, and are borderline for Maths, while TIPI-based estimates remain imprecise—the same qualitative pattern as the main JC-based analysis. The Wave 2→LC specification produces the tightest complementarity estimates (ES = 0.46 for English with SDQ, significantly below one), suggesting that early-adolescent skill interactions are strong predictors of later high-stakes achievement.

Because the design is associational, confounding by unobserved factors remains possible. Oster (2019) bounds for the primary SDQ Focused Behaviour × Cognition interaction indicate that unobservables would need to far exceed the full control set in explanatory power to eliminate the effect ($\delta^* = -6.56$ for Maths, 4.84 for English; Appendix 6.14, Tables 27 and 28).

Standard errors are computed at the individual level rather than clustered at the school level. The GUI Cohort '98 sample was drawn using a two-stage clustered design (Growing Up in Ireland Study Team, 2010), but school identifiers are available only through the Research Microdata File (RMF) pathway; a prior clearance window expired during this work and the re-application process exceeds the current submission timeline.¹³ To gauge the potential magnitude, I inflate the standard errors by factors consistent with the GUI sampling design (assuming roughly 8 sampled students per school and a range of plausible within-school correlations). Even under the most conservative adjustment, the primary SDQ interaction remains significant at the 5% level for both Maths and English (Appendix 6.14, Tables 29 and 30).

Table 13 consolidates all of the robustness evidence for the primary SDQ Focused Behaviour × Cognition result. The interaction coefficient is negative across every linear specification—baseline OLS, baseline-control, IPW-reweighted, z-score outcome, and

¹²LC points are self-reported by the study child in the Wave 4 interview (age 20). Because a 2017 grading reform changed the band structure, a dummy identifies which system applied to each participant and bonus points for Higher Level Maths are adjusted where inconsistently reported.

¹³Published GUI studies employ a range of inference strategies depending on data access: multilevel models (McCoy & Byrne, 2024), survey-weighted GEE (Brennan et al., 2025), within-child fixed effects (Gibbons et al., 2023), and random-effects panel models (McDonnell & Gracia, 2024). Several do not cluster at the school level.

leave-one-out cognition variants-and the translog ES remains below one across sample definitions and shift values. The result survives design-effect-adjusted inference even under conservative clustering assumptions and passes formal FDR screening. No single check reverses the sign or crosses unity.

Table 13: Main-result robustness summary for the primary SDQ Focused Behaviour \times Cognition interaction. Panel A reports the linear interaction coefficient under progressively more demanding specifications. Panel B reports the translog elasticity of substitution (ES) under alternative samples and scaling. Panel C reports inference under conservative clustering assumptions. All SDQ specifications yield complementarity (negative interaction or $ES < 1$); no check reverses the sign or crosses unity.

Specification	Maths	English
<i>Panel A: Linear interaction coefficient</i>		
(1) OLS with controls (Model 3)	-0.057***	-0.074***
(2) + Baseline controls (Model 4, preferred)	-0.052**	-0.063***
(3) IPW-reweighted (selection correction)	-0.074**	-0.100***
(4) Z-score outcome (second-order polynomial)	-0.036***	-0.056***
(5) Cognition leave-one-out range	[-0.047, -0.061] all $p < 0.05$	[-0.062, -0.069] all $p < 0.001$
<i>Panel B: Translog elasticity of substitution (ES)</i>		
(6) Full sample, shift = 0.5 (preferred)	0.482 [0.378, 0.673] $p(ES = 1) = 0.000$	0.464 [0.349, 0.690] $p(ES = 1) = 0.000$
(7) Restricted sample (complete controls)	0.517 [0.370, 0.844]	0.425 [0.325, 0.615]
(8) Alternative shift = 1.0	0.453 [0.350, 0.606]	0.440 [0.337, 0.603]
<i>Panel C: Inference sensitivity</i>		
(9) Design-effect ($\rho = 0.20$, DEFF = 2.40)	$t = -1.99^*$	$t = -2.49^*$
(10) BH-FDR q -value (18 interaction tests)	$q = 0.012^*$	$q = 0.001^{***}$

Notes: Panel A reports Cognition \times SDQ Focused Behaviour interaction coefficients. Row (1): OLS with SES and school controls ($N \approx 3,786$). Row (2): adds Wave 1 age-9 baseline reading, maths, and SDQ subscales. Row (3): inverse-probability-weighted for selection into the restricted sample. Row (4): dependent variable standardised to z -scores; cross-term from second-order polynomial. Row (5): cognition reconstructed by dropping each of three subtests in turn. Panel B: bootstrap percentile 95% CIs from 1,000 replications; $p(ES = 1)$ is a two-sided bootstrap test against Cobb-Douglas. Row (7): translog re-estimated on the restricted OLS sample. Panel C: Row (9) inflates HC1 SEs by $\sqrt{\text{DEFF}}$ under the most conservative intra-class correlation considered ($\rho = 0.20$, average cluster size $\bar{m} = 8$). Row (10): Benjamini-Hochberg adjusted p -values across 18 interaction tests. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

5 Conclusion

This chapter asks a production-function question in an education setting: to what extent can cognitive and noncognitive skills substitute for one another in generating Junior Certificate achievement? Using the Growing Up in Ireland cohort, I estimate linear and translog specifications with predictors measured at age 13 and outcomes observed later in secondary school. The central finding is that the answer depends on how the noncognitive

input is measured - and this measurement-specificity is itself the main contribution.

When the noncognitive input is an SDQ dimension that maps closely to school behaviour, complementarity is stable: the elasticity of substitution is below one in both subjects (approximately 0.48 in Maths and 0.46 in English), the baseline-control linear models show negative cognition-by-Focused Behaviour interaction terms, and these patterns survive value-added controls, inverse-probability reweighting, and alternative functional forms. Read together, the estimates imply that the productivity of one input depends on the level of the other, consistent with the dynamic-skill framework in which later outcomes depend on bundles of interacting capabilities rather than isolated traits (Cunha & Heckman, 2008; Kautz et al., 2014).

When the noncognitive input is instead a TIPI personality trait, the same qualitative pattern weakens and precision falls, especially in subgroup estimates. Part of this contrast is likely psychometric: each TIPI dimension is measured with only two items, yielding internal-consistency reliabilities in the range $\alpha \approx 0.4\text{--}0.6$ (Gosling et al., 2003), compared with five items per SDQ subscale ($\alpha \approx 0.7\text{--}0.8$). Lower reliability inflates sampling variance in the interaction and squared terms that drive the ES calculation, so the wide TIPI confidence intervals may reflect measurement noise as much as genuine construct differences. However, the SDQ-TIPI contrast is not purely a precision story: TIPI point estimates also differ in economically meaningful ways (e.g., the girls-Maths ES above unity), and the Leaving Certificate robustness exercise (Appendix 6.15) shows asymmetric precision across developmental windows as well. The implication is that inferred technology parameters in education depend on whether the noncognitive measure is school-proximal or broad-trait based (Almlund et al., 2011; Humphries & Kosse, 2017), a distinction that matters for how researchers specify and interpret skill inputs.

A further result concerns heterogeneity. The cognitive elasticities are larger in Maths than in English, while the noncognitive elasticities vary less across subjects. Some gender differences in the core parameters are present but moderate, and the pooled interaction tests do not support strong structural breaks by gender in the preferred baseline-control models. This null is itself informative: in the primary SDQ specification, the production technology appears broadly gender-invariant once baseline skill differences are absorbed. The one prominent exception, a TIPI-based girls-Maths estimate above unity, is not stable across measurement choices and is therefore treated as suggestive rather than as a headline finding.

I also address the main credibility risks directly. Including Wave 1 baseline controls preserves the interaction signs and attenuates the magnitudes only modestly. The selection diagnostics and inverse-probability reweighting show limited sensitivity of the primary interactions, and the Oster bounds indicate that a substantial degree of selection on unobservables would be needed to eliminate the focal SDQ interactions. Under a

conservative design-effect inflation motivated by the clustered school sampling, the key interaction terms remain statistically and economically meaningful. No robustness check reverses the sign of the primary complementarity result, and the bounded, discrete nature of OPS outcomes does not drive the translog conclusion: trimming the upper tail of the outcome distribution leaves the primary SDQ ES estimates in the same complementarity range (Appendix 6.14, Table 24).

My interpretations are descriptive. These estimates do not identify treatment effects, and the chapter does not claim that changing one input will causally shift outcomes by the corresponding partial derivative. What the results do provide is a disciplined mapping from observed skill bundles to expected achievement gradients, which is useful for risk profiling, targeting logic, and model-based hypothesis formation in later causal work.

For economics of education, the primary contribution is that measurement choices on the noncognitive margin are consequential for production-function conclusions. Complementarity is empirically stable for attentional regulation but not for generic personality proxies, a distinction that should inform how future studies define and validate skill inputs. The complementarity finding itself - that cognitive and behavioural inputs are jointly productive in an education setting - is the substantive result that measurement-specificity qualifies and strengthens.

The findings point to three concrete directions for further work. First, the measurement-specificity result calls for replication with higher-fidelity noncognitive instruments: multi-item, multi-informant batteries that can separate genuine construct differences from reliability-driven noise in the TIPI contrast. Second, the complementarity result in the Leaving Certificate robustness exercise (Appendix 6.15) suggests that the ES gradient across developmental windows deserves systematic investigation. Do technology parameters shift as cognitive and noncognitive inputs measured at earlier ages predict outcomes at progressively later stages? Third, the descriptive production-function estimates reported here can serve as calibration targets for structural models of skill formation (Cunha et al., 2010) or as priors for targeting designs that allocate intervention intensity based on observed skill bundles. All three directions require either restricted-access school identifiers for clustered inference or linked administrative outcome data, neither of which is available under the current data agreement.

References

- Agostinelli, F., & Wiswall, M. (2024). Estimating the Technology of Children's Skill Formation. *Journal of Political Economy*, *132*(7), 2277–2326.
- Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality Psychology and Economics. In *Handbook of the Economics of Education* (pp. 1–181, Vol. 4). Elsevier.
- Attanasio, O., Meghir, C., & Nix, E. (2020). Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia. *American Economic Review*, *110*(1), 48–85.
- Bertrand, M., & Pan, J. (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics*, *5*(1), 32–64.
- Borghans, L., Meijers, H., & Ter Weel, B. (2008). The Role of Noncognitive Skills in Explaining Cognitive Test Scores. *Economic Inquiry*, *46*(1), 2–12.
- Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement Error in Survey Data. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (pp. 3705–3843, Vol. 5). Elsevier.
- Bowles, S., & Gintis, H. (2002). Schooling in Capitalist America Revisited. *Sociology of Education*, 1–18.
- Brennan, A., et al. (2025). Alcohol Use Among Adolescents in Ireland: A Survey-Weighted Analysis. *BMC Public Health*. <https://doi.org/10.1186/s12889-025-22262-w>
- Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., Barbarin, O. A., Sparling, J. J., & Ramey, C. T. (2012). Adult Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-Up. *Developmental Psychology*, *48*(4), 1033.
- Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental Logarithmic Production Frontiers. *The Review of Economics and Statistics*, 28–45.
- Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, *97*(2), 31–47.
- Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, *43*(4), 738–782.
- Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating The Technology Of COgnitive And Noncognitive Skill Formation. *Econometrica*, *78*(3), 883–931.
- Davis-Kean, P. E. (2005). The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment. *Journal of Family Psychology*, *19*(2), 294.

- Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.
- Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes iq in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.
- Duckworth, A. L., & Seligman, M. E. (2006). Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores. *Journal of Educational Psychology*, *98*(1), 198.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43*(6), 1428–1446.
- Fox, S. E., Levitt, P., & Nelson III, C. A. (2010). How the Timing and Quality of Early Experiences Influence the Development of Brain Architecture. *Child Development*, *81*(1), 28–40.
- Gibbons, A., et al. (2023). Longitudinal Associations Between Screen Time and Wellbeing in Adolescence. *Journal of Youth and Adolescence*, *52*, 517–533. <https://doi.org/10.1007/s10964-022-01721-2>
- Gil-Hernández, C. J. (2021). The (Unequal) Interplay Between Cognitive and Noncognitive Skills in Early Educational Attainment. *American Behavioral Scientist*, *65*(11), 1577–1598.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality*, *37*(6), 504–528.
- Growing Up in Ireland Study Team. (2010). *Growing Up in Ireland: Sample Design and Response in Wave One (9-Year Cohort)* (tech. rep.) (Available at <https://www.ucd.ie/t4cms/GUI-SampleDesign9YearCohort.pdf>). Economic, Social Research Institute, and Trinity College Dublin.
- Hanushek, E. A. (2003). The Failure of Input-Based Schooling Policies. *Economic Journal*, *113*(485), F64–F98.
- Hanushek, E. A., & Woessmann, L. (2012). Do Better Schools Lead To More Growth? Cognitive Skills, Economic Outcomes, And Causation. *Journal of economic growth*, *17*, 267–321.
- Heckman, J. J., & Kautz, T. (2012). Hard Evidence on Soft Skills. *Labour Economics*, *19*(4), 451–464.

- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, *24*(3), 411–482.
- Humphries, J. E., & Kosse, F. (2017). On the Interpretation of Noncognitive Skills—What Is Being Measured and Why It Matters. *Journal of Economic Behavior & Organization*, *136*, 174–185.
- Jackson, C. K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes. *Journal of Political Economy*, *126*(5), 2072–2107.
- Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are Cognitive *g* and Academic Achievement *g* One and the Same *g*? An Exploration on the Woodcock–Johnson and Kaufman Tests. *Intelligence*, *40*(2), 123–138.
- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). Fostering and Measuring Skills: Improving Cognitive and Noncognitive Skills to Promote Lifetime Success.
- Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, Neurobiological, and Behavioral Perspectives on Building America’s Future Workforce. *Proceedings of the National Academy of Sciences*, *103*(27), 10155–10162.
- Lawlor, C. (2019). *An Exploration of Wellbeing Provision in the Junior Cycle Within the Education and Training Board (ETB) Sector in the Context of a Whole School Approach and Guidance Counselling Delivery* [Doctoral dissertation, University of Limerick].
- Lindqvist, E., & Vestman, R. (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101–28.
- McCoy, S., & Byrne, D. (2024). Shadow Education Uptake in Ireland: Inequalities and Wellbeing in a High-Stakes Context. *British Journal of Educational Studies*. <https://doi.org/10.1080/00071005.2024.2331476>
- McDonnell, A., & Gracia, P. (2024). Family Structure and Adolescent Outcomes in Ireland. *Frontiers in Sociology*. <https://doi.org/10.3389/fsoc.2024.1498987>
- National Council for Curriculum and Assessment. (2021). Updated Guidelines for the Primary School Curriculum. https://ncca.ie/media/5062/updated-guidelines-2021_en.pdf
- OECD. (2012). *Public and Private Schools: How Management and Funding Relate to Their Socio-Economic Profile*. <https://doi.org/10.1787/9789264175006-en>
- Olds, D. L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D. W., Henderson Jr, C. R., Hanks, C., Bondy, J., & Holmberg, J. (2004). Effects of Nurse Home-Visiting on Maternal Life Course and Child Development: Age 6 Follow-Up Results of a Randomized Trial. *Pediatrics*, *114*(6), 1550–1559.

- Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204. <https://doi.org/10.1080/07350015.2016.1227711>
- Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The Effects of Single-Sex Compared with Coeducational Schooling on Students' Performance and Attitudes: A Meta-Analysis. *Psychological Bulletin*, *140*(4), 1042.
- Poropat, A. E. (2009). A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance. *Psychological Bulletin*, *135*(2), 322–338.
- Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRPs' Impact on Low-Income Preschoolers' Preacademic Skills: Self-Regulation as a Mediating Mechanism. *Child Development*, *82*(1), 362–378.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. High/Scope Press.
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, *75*(3), 417–453.
- Smyth, E., McCoy, S., & Kingston, G. (2015). Learning from the Evaluation of DEIS. *Dublin: Economic and Social Research Institute*.
- Sofroniou, N., Shiel, G., & Cosgrove, J. (2000). Explaining Performance on the Junior Certificate Examination: A Multilevel Approach. *The Irish Journal of Education/Iris Eireannach an Oideachais*, 25–49.
- Stern, D. I. (2009). Derivation of the Hicks Elasticity of Substitution from the Input Distance Function. *Available at SSRN 1326287*.
- Thiel, H., & Thomsen, S. L. (2013). Noncognitive Skills in Economics: Models, Measurement, and Empirical Evidence. *Research in Economics*, *67*(2), 189–214.
- Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, *113*(485), F3–F33.
- Youth in Mind. (2023). Information for Researchers and Professionals About the Strengths & Difficulties Questionnaires [Accessed: 27/06/2024]. <https://sdqinfo.org/>

6 Appendices

The following appendices provide supplementary material, including psychometric tool descriptions, production-function derivations, extended analyses, and robustness checks that support the core findings of this chapter.

1. Extended literature review.
2. TIPI and SDQ instruments (scoring and sample items).
3. Cobb-Douglas production function: two-input and three-input derivations (marginal products, output elasticities, ES, MRTS).
4. MRTS analysis by gender and subject.
5. Translog production function derivations.
6. CES production function (two-input and three-input).
7. Comparison of CES model variants.
8. Interpretive implications.
9. Empirical considerations.
10. Selection and attrition diagnostics.
11. Additional robustness checks.
12. Complementarity across developmental windows.

6.1 Extended literature review

6.1.1 Noncognitive skills in the achievement literature

Noncognitive skills rival cognitive ability in predicting educational attainment, labour market success, and health (Heckman et al., 2006; Kautz et al., 2014). Conscientiousness and self-discipline are particularly strong predictors of grades (Almlund et al., 2011; Duckworth & Seligman, 2006), and noncognitive skills may be especially important at the lower end of the earnings distribution (Lindqvist & Vestman, 2011). The interaction between cognitive and noncognitive skills operates through both effort and intrinsic motivation channels (Borghans et al., 2008), while skill development follows a dynamic process in which early investments raise the productivity of later ones (Cunha & Heckman, 2008; Kautz et al., 2014).

6.1.2 Gender differences and measurement challenges

Boys are more susceptible to behavioural problems, especially in disadvantaged environments (Bertrand & Pan, 2013), and girls tend to outperform boys in self-discipline even at similar IQ levels (Duckworth & Seligman, 2006). Noncognitive skills are more malleable than cognitive skills during adolescence (Kautz et al., 2014), creating a window for targeted intervention.

Measurement remains challenging. Definitions vary across disciplines: labour economists treat noncognitive skills as a second dimension of heterogeneity, while psychologists use psychometric personality constructs (Humphries & Kosse, 2017; Thiel & Thomsen, 2013). No single global measure exists; this paper uses two instruments - SDQ and TIPI - that capture different facets of noncognitive skills, allowing direct comparison of how measurement choice affects production-function estimates.

6.2 TIPI questionnaire

The TIPI (Gosling et al., 2003) measures the Big Five personality traits-Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience-using ten items (two per trait) on a 7-point Likert scale. Its brevity makes it practical for large-scale surveys such as GUI, though reliability and validity are lower than for full-length inventories. In this study, primary caregivers rated the study child; scoring involves reverse-coding one item per trait and averaging the pair.

Disagree strongly	Disagree moderately	Disagree a little	Neither disagree nor agree	Agree a little	Agree moderately	Agree strongly
1	2	3	4	5	6	7

Characteristic	Grade (1-7)
1. Extraverted, enthusiastic	_____
2. Critical, quarrelsome	_____
3. Dependable, self-disciplined	_____
4. Anxious, easily upset	_____
5. Open to new experiences, complex	_____
6. Reserved, quiet	_____
7. Sympathetic, warm	_____
8. Disorganized, careless	_____
9. Calm, emotionally stable	_____
10. Conventional, uncreative	_____

Conscientiousness	Agreeableness	Emotional Stability
Score for #3: _____	Score for #7: _____	Score for #9: _____
+ (8 - Score for #8): _____	+ (8 - Score for #2): _____	+ (8 - Score for #4): _____
_____	_____	_____
= _____	= _____	= _____
Divide your answer by 2.	Divide your answer by 2.	Divide your answer by 2.
Conscientiousness Score	Agreeableness Score =	Emotional Stability Score
= _____	_____	= _____

Openness to Experience	Extraversion
Score for #5: _____	Score for #1: _____
+ (8 - Score for #10): _____	+ (8 - Score for #6): _____
= _____	= _____
Divide your answer by 2.	Divide your answer by 2.
Openness to Experience Score =	Extraversion Score = _____

6.3 SDQ questionnaire

The Strengths and Difficulties Questionnaire (Goodman, 1997) is a 25-item behavioural screening tool for children aged 3–16, comprising five subscales: Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer-relationship Problems, and Prosocial Behaviour. Items are scored on a 3-point scale (“Not True”, “Somewhat True”, “Certainly True”). In this study, the primary caregiver completed the parent version; the Prosocial subscale was excluded because its scoring in the anonymised microdata files was inconsistent with the other four subscales. Higher difficulty subscale scores indicate greater problems (range 0–10 per subscale).

Sample items (from Youth in Mind (2023), <https://sdqinfo.org/>):

Strengths and Difficulties Questionnaire

For each item, please mark the box for Not True, Somewhat True or Certainly True. Answer all items as best you can even if you are not absolutely certain. Please give your answers on the basis of the child’s behaviour over the last six months.

	Not True	Somewhat True	Certainly True
1. Considerate of other people's feelings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Restless, overactive, cannot stay still for long	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Often complains of headaches, stomach-aches or sickness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Shares readily with other children (treats, toys, pencils etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Often has temper tantrums or hot tempers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Rather solitary, tends to play alone	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Generally obedient, usually does what adults request	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Many worries, often seems worried	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Helpful if someone is hurt, upset or feeling ill	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Constantly fidgeting or squirming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Has at least one good friend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Often fights with other children or bullies them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. Often unhappy, down-hearted or tearful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. Generally liked by other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. Easily distracted, concentration wanders	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Nervous or clingy in new situations, easily loses confidence	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Kind to younger children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. Often lies or cheats	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Picked on or bullied by other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Often volunteers to help others (parents, teachers, other children)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. Thinks things out before acting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. Steals from home, school or elsewhere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. Gets on better with adults than with other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. Many fears, easily scared	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. Sees tasks through to the end, good attention span	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The SDQ is divided into five sections, each containing five questions:

1. Emotional Symptoms Scale:

- 3. Often complains of headaches, stomach-aches or sickness
- 8. Many worries, often seems worried

- 13. Often unhappy, down-hearted or tearful
- 16. Nervous or clingy in new situations, easily loses confidence
- 24. Many fears, easily scared

2. Conduct Problems Scale:

- 5. Often has temper tantrums or hot tempers
- 7. Generally obedient, usually does what adults request (reverse scored)
- 12. Often fights with other children or bullies them
- 18. Often lies or cheats
- 22. Steals from home, school or elsewhere

3. Hyperactivity Scale:

- 2. Restless, overactive, cannot stay still for long
- 10. Constantly fidgeting or squirming
- 15. Easily distracted, concentration wanders
- 21. Thinks things out before acting (reverse scored)
- 25. Sees tasks through to the end, good attention span (reverse scored)

4. Peer Problems Scale:

- 6. Rather solitary, tends to play alone
- 11. Has at least one good friend (reverse scored)
- 14. Generally liked by other children (reverse scored)
- 19. Picked on or bullied by other children
- 23. Gets on better with adults than with other children

5. Prosocial Scale:

- 1. Considerate of other people's feelings
- 4. Shares readily with other children (treats, toys, pencils etc.)
- 9. Helpful if someone is hurt, upset or feeling ill
- 17. Kind to younger children
- 20. Often volunteers to help others (parents, teachers, other children)

Note: Items marked as "reverse scored" are phrased positively, so their scores are reversed when calculating the total for that scale.

6.4 Cobb-Douglas With Two Inputs

This section presents the two-input Cobb-Douglas production function used to model the relationship between cognitive and noncognitive skills and academic achievement. The functional form is:

$$Y = AC^\alpha N^\beta \quad (4)$$

Where:

Y : Academic achievement (output)

A : Total factor productivity (scaling factor)

C : Cognitive skill input

N : Noncognitive skill input

α, β : Output elasticities of each input

The exponents α and β capture the percentage change in academic performance resulting from a 1% change in cognitive and noncognitive skills, respectively. This formulation assumes constant elasticity of substitution ($\sigma = 1$) and exhibits decreasing, constant, or increasing returns to scale depending on the sum of $\alpha + \beta$.

Although a more flexible three-input version is examined in the next section, this simplified model remains informative. The similarity in parameters between the two- and three-input specifications shows that a two-factor structure - focused on cognition and a single noncognitive measure - captures most of the relevant variance in achievement outcomes, supporting its standalone presentation.

6.4.1 Marginal Products (MPs)

Marginal products represent the additional output generated by a one-unit increase in a given input while holding the other input constant. For the Cobb-Douglas production function:

$$Y = AC^\alpha N^\beta$$

The marginal product of cognition (C) is:

$$f_C = \left. \frac{\partial Y}{\partial C} \right|_{N=N_0} = A\alpha C^{\alpha-1} (N_0)^\beta = A\alpha \frac{C^\alpha N^\beta}{C} = \alpha \frac{Y}{C}$$

The marginal product of noncognitive skills (N) is:

$$f_N = \frac{\partial Y}{\partial N} \Big|_{C=C_0} = A\beta C_0^\alpha N^{\beta-1} = A\beta \frac{C^\alpha N^\beta}{N} = \beta \frac{Y}{N}$$

To assess whether marginal returns increase or decrease, we take the second derivatives:

$$\frac{\partial f_C}{\partial C} = \frac{\partial^2 Y}{\partial C^2} = A\alpha(\alpha - 1)C^{\alpha-2}N^\beta = \alpha(\alpha - 1)\frac{Y}{C^2}$$

$$\frac{\partial f_N}{\partial N} = \frac{\partial^2 Y}{\partial N^2} = A\beta(\beta - 1)N^{\beta-2}C^\alpha = \beta(\beta - 1)\frac{Y}{N^2}$$

These second derivatives help determine whether each input exhibits diminishing returns, which occurs when the expressions are negative (i.e., when $\alpha < 1$ or $\beta < 1$). In this study, estimated values of α and β are generally below 1, implying decreasing marginal returns to both cognition and noncognitive skills in the production of academic achievement.

6.4.2 Output elasticities (OEs)

Parameters estimation:

$$\ln(Y) = \ln(A) + \alpha \ln(C) + \beta \ln(N) \quad (5)$$

With output elasticities defined as:

$$OE_C = \frac{\partial \ln(Y)}{\partial \ln(C)} \Big|_{N=N_0} = \alpha \quad (6)$$

$$OE_N = \frac{\partial \ln(Y)}{\partial \ln(N)} \Big|_{C=C_0} = \beta \quad (7)$$

If we define the scale elasticity (SCE) as the scale change, often measured by the percent change in output from a simultaneous 1% change in all inputs, then: $OE_C + OE_N = \alpha + \beta$.

6.4.3 elasticity of substitution

$$\sigma = \frac{\frac{d(C/N)}{(C/N)}}{\frac{d(MP_C/MP_N)}{(MP_C/MP_N)}} \quad (8)$$

Where $MP_C(f_C)$ and $MP_N(f_N)$ are the marginal products of C and N respectively. For the Cobb-Douglas production function, the σ between the two inputs is always equal to 1. This is a key property of the Cobb-Douglas function. To illustrate this for the two-input case:

$$Y = AC^\alpha N^{1-\alpha} \quad (9)$$

The marginal products are:

$$f_C = \frac{\partial Y}{\partial C} = \alpha AC^{\alpha-1} N^{1-\alpha} \quad (10)$$

$$f_N = \frac{\partial Y}{\partial N} = (1 - \alpha) AC^\alpha N^{-\alpha} \quad (11)$$

The ratio of marginal products is:

$$\frac{f_C}{f_N} = \frac{\alpha N}{(1 - \alpha)C} \quad (12)$$

If we calculate σ , we find:

$$\sigma = \frac{\frac{d(C/N)}{(C/N)}}{\frac{d((\alpha N)/((1-\alpha)C))}{(\alpha N)/((1-\alpha)C)}} = 1 \quad (13)$$

This unit value is a structural property of the two-input Cobb-Douglas form, which imposes constant unitary elasticity of substitution between cognition and noncognition by construction.

6.4.4 Estimation and discussion

The cognitive factor α , representing the output elasticity of cognition in this Cobb-Douglas specification, is the most significant predictor of academic performance across all models. For Maths (Table 14), boys show slightly higher cognitive output elasticities ($\alpha = 0.808$ for TIPI, 0.791 for SDQ) compared to girls ($\alpha = 0.776$ for TIPI, 0.751 for SDQ). This pattern is mirrored in English (Table 15), with boys' α ranging from 0.477 to 0.492 and girls' from 0.419 to 0.435. The full sample results fall between these gender-specific values, as expected.

The noncognitive factor β , representing the output elasticity of the noncognitive skill in this specification, is smaller in magnitude but significant across all models, as expected. Focused Behaviour consistently exhibits stronger associations than Conscientiousness. For example, in the full sample Maths model, $\beta = 0.038$ for the SDQ model while $\beta = 0.024$ for the TIPI model.

Boys consistently demonstrate a slightly stronger cognitive component in both subjects. However, girls show stronger effects of noncognitive factors, particularly in Maths. This is especially evident with the SDQ measures, where girls' OE_N for Maths is 0.051 compared to boys' 0.029.

Both cognitive and noncognitive factors appear to have a stronger influence on Maths performance compared to English. This is evident in the higher values of both α and β for Maths across all models. For example, in the full sample SDQ model, $\alpha_{Maths} = 0.761$ while $\alpha_{English} = 0.424$.

Table 14: Two-input Cobb-Douglas estimates for Junior Certificate Maths achievement by noncognitive measure set (TIPI and SDQ) and subgroup (full sample, boys, girls). Reported quantities include estimated elasticities, implied marginal products, and returns to scale; values are descriptive model summaries for the estimation sample.

Parameter	Maths			
	Full Sample	Boys	Girls	Δ (Boys - Girls)
TIPI Model				
A	0.256*** (0.015)	0.224*** (0.019)	0.263*** (0.021)	-0.039
α (Cognition)	0.781*** (0.012)	0.808*** (0.018)	0.776*** (0.018)	0.032
β (Conscientiousness)	0.024*** (0.003)	0.022*** (0.004)	0.024*** (0.004)	-0.002
Marginal Product (Cognition)	0.075	0.077	0.076	0.001
Marginal Product (Conscientiousness)	0.053	0.051	0.051	0.000
Returns to Scale	0.805	0.830	0.800	0.030
SDQ Model				
A	0.268*** (0.018)	0.236*** (0.025)	0.276*** (0.024)	-0.040
α (Cognition)	0.761*** (0.015)	0.791*** (0.023)	0.751*** (0.020)	0.040
β (Focused Behaviour)	0.038*** (0.006)	0.029*** (0.008)	0.051*** (0.010)	-0.022
Marginal Product (Cognition)	0.073	0.075	0.073	0.002
Marginal Product (Focused Behaviour)	0.049	0.039	0.061	-0.022
Returns to Scale	0.800	0.820	0.801	0.019
Observations	5,631	2,667	2,801	

*Standard errors in parentheses. Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: The table displays estimates for the Cobb-Douglas production function applied to Maths scores using two inputs. The TIPI Model focuses on Cognition and Conscientiousness, while the SDQ Model considers Cognition and Focused Behaviour. Observations represent the number of data points for each group.

The marginal products further support these findings. Across all models, the marginal product for cognition is always higher than for noncognitive factors, reinforcing the dominant role of cognitive abilities in academic achievement.

The sum of α and β in all models is less than 1, which indicates decreasing returns to scale in the production of academic achievement. This implies that proportional increases in both cognitive and noncognitive inputs would result in less than proportional increases in academic output. For example, in the full sample Maths SDQ model, $\alpha + \beta = 0.761 + 0.038 = 0.800 < 1$. The educational production process appears more efficient for Maths, as evidenced by higher returns to scale compared to English.

Table 15: Two-input Cobb-Douglas estimates for Junior Certificate English achievement by noncognitive measure set (TIPI and SDQ) and subgroup (full sample, boys, girls). Reported quantities include estimated elasticities, implied marginal products, and returns to scale; values are descriptive model summaries for the estimation sample.

Parameter	English			
	Full Sample	Boys	Girls	Δ (Boys - Girls)
TIPI Model				
A	1.313*** (0.060)	1.017*** (0.072)	1.388*** (0.084)	-0.371
α (Cognition)	0.441*** (0.010)	0.492*** (0.015)	0.435*** (0.013)	0.057
β (Conscientiousness)	0.014*** (0.002)	0.010** (0.003)	0.010** (0.003)	0.000
Marginal Product (Cognition)	0.045	0.048	0.046	0.002
Marginal Product (Conscientiousness)	0.032	0.025	0.023	0.002
Returns to Scale	0.454	0.502	0.445	0.057
SDQ Model				
A	1.368*** (0.079)	1.066*** (0.107)	1.434*** (0.105)	-0.368
α (Cognition)	0.424*** (0.013)	0.477*** (0.022)	0.419*** (0.017)	0.058
β (Focused Behaviour)	0.027*** (0.005)	0.019*** (0.008)	0.026*** (0.008)	-0.007
Marginal Product (Cognition)	0.043	0.047	0.044	0.003
Marginal Product (Focused Behaviour)	0.037	0.027	0.034	-0.007
Returns to Scale	0.452	0.496	0.445	0.051
Observations	5,631	2,667	2,801	

*Standard errors in parentheses. Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: The table displays estimates for the Cobb-Douglas production function applied to English scores using two inputs. The TIPI Model focuses on Cognition and Conscientiousness, while the SDQ Model considers Cognition and Focused Behaviour. Observations represent the number of data points for each group.

Overall, the two-input Cobb-Douglas model offers a parsimonious yet informative perspective on how key skill inputs shape academic performance.

6.5 Cobb-Douglas With Three Inputs

We begin with a Cobb-Douglas production function incorporating three inputs:

$$Y = f(C, N_E, N_I) = AC^\alpha N_E^{\beta_1} N_I^{\beta_2} \quad (14)$$

Where:

Y : Total output/Grade function/Academic achievement

A : Total factor productivity/scaling factor

C : Input representing cognition

N_E, N_I : Inputs representing noncognitive measures

α, β_1, β_2 : Exponents determining the output response to each input

This function assumes a Cobb-Douglas form, where the exponents α , β_1 , and β_2 represent the output elasticities - capturing the proportional impact of each input on academic achievement. C is a measure of cognition and N_E and N_I are noncognitive measures. N_E captures emotional traits (Emotional Resilience for SDQ and Emotional Stability for TIPI), while N_I captures behavioural traits (Focused Behaviour for SDQ and Conscientiousness for TIPI). These were selected based on their consistent significance in prior regressions and relatively high pairwise correlations (0.407 for SDQ and 0.409 for TIPI). Correlation analyses confirmed their centrality, with these pairs exhibiting the highest coefficients among all subscales - supporting their use as representative noncognitive dimensions. The use of separate noncognitive inputs allows us to better capture the multidimensional nature of noncognitive skills and their potentially different impacts on academic achievement. While other subscales were available, these four stood out as the most relevant for predicting academic performance in both Maths and English.

6.5.1 Marginal products (MPs)

Marginal products (MPs) represent the change in total output resulting from a one-unit increase in a specific input while holding all other inputs constant.

The marginal products for each input represent the change in academic achievement resulting from a one-unit increase in the respective input, holding the others constant.

$$f_C = \left. \frac{\partial f}{\partial C} \right|_{N_E=N_{E0}, N_I=N_{I0}} = A\alpha C^{\alpha-1} (N_{E0})^{\beta_1} (N_{I0})^{\beta_2} \quad (15)$$

$$f_{N_E} = \left. \frac{\partial f}{\partial N_E} \right|_{C=C_0, N_I=N_{I0}} = AC_0^\alpha \beta_1 (N_E)^{\beta_1-1} (N_{I0})^{\beta_2} \quad (16)$$

$$f_{N_I} = \left. \frac{\partial f}{\partial N_I} \right|_{C=C_0, N_E=N_{E0}} = AC_0^\alpha (N_{E0})^{\beta_1} \beta_2 (N_I)^{\beta_2-1} \quad (17)$$

6.5.2 Output elasticities (OEs)

Output elasticities measure the responsiveness of output to a change in an input, expressed in percentage terms.

Given:

$$\ln(Y) = \ln(A) + \alpha \ln(C) + \beta_1 \ln(N_E) + \beta_2 \ln(N_I) \quad (18)$$

With output elasticities defined as:

$$OE_C = \left. \frac{\partial \ln(Y)}{\partial \ln(C)} \right|_{N_E=N_{E0}, N_I=N_{I0}} = \alpha \quad (19)$$

$$OE_{N_E} = \left. \frac{\partial \ln(Y)}{\partial \ln(N_E)} \right|_{C=C_0, N_I=N_{I0}} = \beta_1 \quad (20)$$

$$OE_{N_I} = \left. \frac{\partial \ln(Y)}{\partial \ln(N_I)} \right|_{C=C_0, N_E=N_{E0}} = \beta_2 \quad (21)$$

The output elasticities for cognition and each noncognitive input are equal to their respective exponents α , β_1 , and β_2 , as is standard in log-linear Cobb-Douglas models.

The scale elasticity (SCE) measures by the percent change in output from a simultaneous 1% change in all inputs, then:

$$SCE = OE_C + OE_{N_E} + OE_{N_I} = \alpha + \beta_1 + \beta_2 \quad (22)$$

6.5.3 elasticity of substitution

The elasticity of substitution (ES, σ) is defined as the degree to which the marginal rate of substitution between two inputs varies as the ratio of the quantity of those inputs varies while output is held constant (Stern, 2009):

$$\sigma = \frac{\frac{d(X/Y)}{(X/Y)}}{\frac{d(MP_X/MP_Y)}{(MP_X/MP_Y)}} \quad (23)$$

Where $MP_X(f_X)$ and $MP_Y(f_Y)$ are the marginal products of X and Y respectively. For the Cobb-Douglas production function, the σ between any two inputs is always = 1. This implies that inputs are neither strong substitutes nor strong complements; rather, they are unitary substitutes, meaning a 1% increase in one input requires a 1% decrease in another to keep output constant. To illustrate this for the three-input case:

$$\frac{f_C}{f_{N_E}} = \frac{\alpha f/C}{\beta_1 f/N_E} = \frac{\alpha N_E}{\beta_1 C} \quad (24)$$

If we were to calculate σ_{C,N_E} , we would find:

$$\sigma = \frac{\frac{d(C/N_E)}{(C/N_E)}}{\frac{(\alpha N_E)/(\beta_1 C)}{(\alpha N_E)/(\beta_1 C)}} = 1 \quad (25)$$

6.5.4 Estimation and discussion

To extend the two-input setup, I estimate three-input Cobb-Douglas models in which cognition enters jointly with two noncognitive margins. For TIPI specifications:

$$JC_{M,E} = A(\text{Cognition})^\alpha (\text{EmotionalStability})^{\beta_1} (\text{Conscientiousness})^{\beta_2} \quad (26)$$

For SDQ specifications:

$$JC_{M,E} = A(\text{Cognition})^\alpha (\text{EmotionalResilience})^{\beta_1} (\text{FocusedBehaviour})^{\beta_2} \quad (27)$$

where $JC_{M,E}$ denotes Junior Certificate scores in Maths (M) and English (E).

Across all models, cognition continues to be the strongest and most consistent predictor of academic achievement. However, this expanded analysis reveals subtle gender differences, with boys exhibiting marginally higher cognitive elasticities in both subjects.

Two distinct noncognitive factors enhance analytical depth. Though smaller in effect, noncognitive traits preserve statistical significance, particularly within SDQ models, which validates their role in educational outcomes. Girls exhibit stronger noncognitive effects - particularly in Maths - underscoring the importance of behavioural traits in female academic success and challenging one-dimensional gender narratives.

Subject-wise comparisons indicate that cognition exerts a more substantial influence on Maths than on English across all models. Noncognitive factors, particularly for girls, contribute more significantly to Maths performance than might be expected.

These findings reinforce the greater predictive validity of the SDQ over the TIPI for academic outcomes, likely due to its context-specific behavioural focus. The examination of marginal products reinforces cognition's dominant role while also shedding light at the non-trivial contributions of noncognitive factors. The observed decreasing returns to scale, more pronounced in English, imply that proportional increases in all inputs yield diminishing academic gains.

6.6 Marginal Rate of Technical Substitution (MRTS) for Cobb-Douglas Production Functions with Two and Three Inputs

The Marginal Rate of Technical Substitution (MRTS) originates in production theory and is applied here to educational achievement. In this context, the MRTS represents the rate

Table 16: Three-input Cobb-Douglas estimates for Junior Certificate Maths achievement using cognition and two noncognitive inputs (TIPI: emotional stability and conscientiousness; SDQ: emotional resilience and focused behaviour), by subgroup (full sample, boys, girls). Estimates are descriptive and reported with standard errors in parentheses.

Parameter	Full Sample	Boys	Girls	Δ (Boys - Girls)
TIPI Model				
A	0.257*** (0.015)	0.225*** (0.019)	0.264*** (0.021)	-0.039
α (Cognition)	0.777*** (0.012)	0.804*** (0.018)	0.774*** (0.018)	0.030
β_1 (Emotional Stability)	0.009** (0.003)	0.013** (0.005)	0.005 (0.004)	0.008
β_2 (Conscientiousness)	0.022*** (0.003)	0.020*** (0.004)	0.023*** (0.004)	-0.003
Marginal Product (Cognition)	0.075	0.077	0.076	0.001
Marginal Product (Emotional Stability)	0.021	0.028	0.012	0.016
Marginal Product (Conscientiousness)	0.050	0.046	0.049	-0.004
Returns to Scale	0.809	0.837	0.802	0.034
SDQ Model				
A	0.264*** (0.018)	0.228*** (0.025)	0.273*** (0.024)	-0.045
α (Cognition)	0.753*** (0.015)	0.784*** (0.023)	0.744*** (0.020)	0.040
β_1 (Emotional Resilience)	0.027** (0.008)	0.035* (0.013)	0.023* (0.010)	0.012
β_2 (Focused Behaviour)	0.036*** (0.006)	0.027*** (0.008)	0.047*** (0.010)	-0.020
Marginal Product (Cognition)	0.073	0.075	0.073	0.002
Marginal Product (Emotional Resilience)	0.031	0.040	0.027	0.013
Marginal Product (Focused Behaviour)	0.046	0.036	0.057	-0.021
Returns to Scale	0.816	0.846	0.814	0.032
Observations	5,631	2,667	2,801	

Standard errors in parentheses. Signif. Codes: ***: 0.001, **: 0.01, *: 0.05

Table 17: Three-input Cobb-Douglas estimates for Junior Certificate English achievement using cognition and two noncognitive inputs (TIPI: emotional stability and conscientiousness; SDQ: emotional resilience and focused behaviour), by subgroup (full sample, boys, girls). Estimates are descriptive and reported with standard errors in parentheses.

Parameter	Full Sample	Boys	Girls	Δ (Boys - Girls)
TIPI Model				
A	1.316*** (0.060)	1.019*** (0.072)	1.390*** (0.084)	-0.371
α (Cognition)	0.439*** (0.010)	0.490*** (0.015)	0.434*** (0.013)	0.056
β_1 (Emotional Stability)	0.004 (0.003)	0.007 (0.004)	0.003 (0.003)	0.004
β_2 (Conscientiousness)	0.013*** (0.002)	0.009** (0.003)	0.010** (0.003)	-0.001
Marginal Product (Cognition)	0.045	0.048	0.046	0.002
Marginal Product (Emotional Stability)	0.008	0.015	0.006	0.009
Marginal Product (Conscientiousness)	0.031	0.022	0.022	0.000
Returns to Scale	0.456	0.506	0.446	0.060
SDQ Model				
A	1.365*** (0.082)	1.050*** (0.107)	1.433*** (0.105)	-0.383
α (Cognition)	0.423*** (0.013)	0.473*** (0.022)	0.418*** (0.017)	0.055
β_1 (Emotional Resilience)	0.005 (0.007)	0.016 (0.012)	0.004 (0.008)	0.012
β_2 (Focused Behaviour)	0.027*** (0.005)	0.018*** (0.008)	0.025*** (0.008)	-0.007
Marginal Product (Cognition)	0.043	0.047	0.044	0.003
Marginal Product (Emotional Resilience)	0.006	0.019	0.005	0.014
Marginal Product (Focused Behaviour)	0.036	0.025	0.033	-0.008
Returns to Scale	0.454	0.507	0.447	0.060
Observations	5,631	2,667	2,801	

Standard errors in parentheses. Signif. Codes: ***: 0.001, **: 0.01, *: 0.05

at which one input (e.g., cognition) can be substituted for another (e.g., noncognitive skills) while maintaining the same level of output (academic performance). Mathematically, the MRTS is defined as the negative slope of the isoquant curve in input space.

In the Cobb-Douglas production function models, the MRTS helps us quantify the trade-offs between cognitive and noncognitive inputs in educational achievement. Specifically, we would be able to answer questions such as: how much improvement in noncognitive skills is needed to compensate for a deficit in cognitive abilities? To what extent can enhancements in one type of skill make up for deficiencies in another? Do these trade-offs differ between subjects (Maths vs. English) or between genders?

These questions are examined using both two-input and three-input Cobb-Douglas production functions, incorporating TIPI and SDQ measures of noncognitive skills (Conscientiousness and Emotional Stability for the TIPI, and Focused Behaviour and Emotional Resilience for the SDQ).

6.6.1 Definition - Three Inputs

We first start with a Cobb-Douglas production function with three-inputs:

$$Y = AC^\alpha N_E^{\beta_1} N_I^{\beta_2} \quad (28)$$

Given:

$$f_C = \frac{\partial Q}{\partial C} = \alpha AC^{\alpha-1} N_E^{\beta_1} N_I^{\beta_2} \quad (29)$$

$$f_{NE} = \frac{\partial Q}{\partial N_E} = \beta_1 AC^\alpha N_E^{\beta_1-1} N_I^{\beta_2} \quad (30)$$

$$f_{NI} = \frac{\partial Q}{\partial N_I} = \beta_2 AC^\alpha N_E^{\beta_1} N_I^{\beta_2-1} \quad (31)$$

Then MRTS:

$$MRTS_{N_E,C} = -\frac{dN_E}{dC} = \frac{f_C}{MP_{NE}} = \frac{\alpha AC^{\alpha-1} N_E^{\beta_1} N_I^{\beta_2}}{\beta_1 AC^\alpha N_E^{\beta_1-1} N_I^{\beta_2}} = \frac{\alpha}{\beta_1} \frac{N_E}{C} \quad (32)$$

$$MRTS_{N_I,C} = -\frac{dN_I}{dC} = \frac{f_C}{f_{NI}} = \frac{\alpha AC^{\alpha-1} N_E^{\beta_1} N_I^{\beta_2}}{\beta_2 AC^\alpha N_E^{\beta_1} N_I^{\beta_2-1}} = \frac{\alpha}{\beta_2} \frac{N_I}{C} \quad (33)$$

$$MRTS_{N_I,N_E} = -\frac{dN_I}{dN_E} = \frac{f_{NE}}{f_{NI}} = \frac{\beta_1 AC^\alpha N_E^{\beta_1-1} N_I^{\beta_2}}{\beta_2 AC^\alpha N_E^{\beta_1} N_I^{\beta_2-1}} = \frac{\beta_1}{\beta_2} \frac{N_I}{N_E} \quad (34)$$

6.6.2 Definition - Two Inputs

Building on the MRTS formulation for the three-input case, we now define the MRTS for a two-input Cobb-Douglas function.

Given:

$$Y = AC^\alpha N^\beta \quad (35)$$

Marginal products:

$$f_C = \frac{\partial Y}{\partial C} \Big|_{N = N_0} = A\alpha C^{\alpha-1} (N_0)^\beta \quad (36)$$

$$f_N = \frac{\partial Y}{\partial N} \Big|_{C = C_0} = A\beta C_0^\alpha (N)^\beta \quad (37)$$

$$MRTS_{N,C} = \frac{f_C}{f_N} = \frac{A\alpha C^{\alpha-1} N^\beta}{A\beta C^\alpha N^{\beta-1}} = \frac{\alpha N}{\beta C} \quad (38)$$

$MRTS_{N,C}$ is:

$$MRTS_{N,C} = \frac{\alpha N}{\beta C} \quad (39)$$

$MRTS_{N,C}$ represents how much the noncognitive input (N) needs to increase to compensate for a unit decrease in cognition (C) while maintaining the same level of output (Y).

6.6.3 Estimation and Discussion

Across most specifications, MRTS values for noncognitive skill relative to cognition exceed 1, which means more than one unit of noncognitive input is needed to offset a one-unit decline in cognition at constant output. This pattern reinforces the core empirical result that cognition is generally the harder input to replace in the achievement production process.

The subject comparison is consistent with this interpretation. In TIPI models, Conscientiousness-for-Cognition MRTS is slightly higher in Maths than in English for the full sample (1.420 vs 1.390). In SDQ models, Focused Behaviour-for-Cognition MRTS is likewise higher in Maths than in English (1.504 vs 1.176), indicating relatively stronger cognitive dominance in Maths. Reciprocals are usually below 1, which implies that cognition compensates for noncognitive shortfalls more easily than noncognitive inputs compensate for cognitive shortfalls.

Gender differences are present but uneven. In several specifications, boys require more noncognitive input to substitute for cognition than girls, while English shows larger subgroup variation overall, especially under TIPI. By contrast, SDQ-based MRTS

Model	MRTS Type	Marginal Rates of Technical Substitution (MRTS)			
		Full Sample	Girls	Boys	Δ (Boys - Girls)
Maths (TIPI)	Emotional Stability for Cognition	3.600	6.255	2.724	-3.531
	Conscientiousness for Cognition	1.511	1.532	1.674	0.142
	Conscientiousness for Emo. Stability	0.420	0.245	0.615	0.370
Maths (SDQ)	Emotional Resilience for Cognition	2.322	2.643	1.847	-0.796
	Focused Behaviour for Cognition	1.599	1.272	2.094	0.822
	Focused Behaviour for Emo. Resilience	0.689	0.481	1.134	0.653
English (TIPI)	Emotional Stability for Cognition	5.427	7.358	3.190	-4.168
	Conscientiousness for Cognition	1.444	2.054	2.161	0.107
	Conscientiousness for Emo. Stability	0.266	0.279	0.677	0.398
English (SDQ)	Emotional Resilience for Cognition	7.623	8.552	2.516	-6.036
	Focused Behaviour for Cognition	1.191	1.316	1.841	0.525
	Focused Behaviour for Emo. Resilience	0.156	0.154	0.732	0.578

Note: The table presents the Marginal Rates of Technical Substitution for 3-input Cobb-Douglas models in both Maths and English. The TIPI model uses Emotional Stability, Conscientiousness, and Cognition as inputs, while the SDQ model utilizes Emotional Resilience, Focused Behaviour, and Cognition. MRTS indicates the rate at which one input can be substituted for another while maintaining the same level of output. Δ represents the difference in MRTS between Boys and Girls.

Table 18: Marginal rates of technical substitution (MRTS) implied by three-input Cobb-Douglas models for Maths and English, by noncognitive measure set (TIPI and SDQ) and subgroup (full sample, boys, girls). MRTS values indicate within-model input trade-offs at observed sample levels and are descriptive model implications.

Model	MRTS Type	Marginal Rates of Technical Substitution (MRTS)			
		Full Sample	Girls	Boys	Δ (Boys - Girls)
Maths (TIPI)	Conscientiousness for Cognition	1.420	1.475	1.520	0.045
	Cognition for Conscientiousness	0.704	0.678	0.658	-0.020
Maths (SDQ)	Focused Behaviour for Cognition	1.504	1.192	1.925	0.733
	Cognition for Focused Behaviour	0.665	0.839	0.519	-0.320
English (TIPI)	Conscientiousness for Cognition	1.390	1.966	1.950	-0.016
	Cognition for Conscientiousness	0.720	0.509	0.513	0.004
English (SDQ)	Focused Behaviour for Cognition	1.176	1.292	1.744	0.452
	Cognition for Focused Behaviour	0.850	0.774	0.573	-0.201

Note: The table presents the Marginal Rates of Technical Substitution for 2-input Cobb-Douglas models in both Maths and English. The TIPI model uses Cognition and Conscientiousness as inputs, while the SDQ model utilizes Cognition and Focused Behaviour. MRTS is calculated as the ratio of the marginal product of one input to the marginal product of the other, indicating how inputs can be substituted while maintaining the same level of output.

Table 19: Marginal rates of technical substitution (MRTS) implied by two-input Cobb-Douglas models for Maths and English, by noncognitive measure set (TIPI and SDQ) and subgroup (full sample, boys, girls). MRTS values indicate within-model input trade-offs at observed sample levels and are descriptive model implications.

values are more stable across subjects and subgroups, which is consistent with the broader finding that school-proximal behavioural measures map more consistently into achievement

trade-offs than broader personality proxies.

6.7 Derivations for a Translog Production Function with Two Inputs

$$Y = AC^\alpha N^\beta \exp \left\{ \frac{1}{2} \gamma_1 [\ln(C)]^2 + \frac{1}{2} \gamma_2 [\ln(N)]^2 + \gamma_{12} \ln(C) \ln(N) \right\} \quad (40)$$

Where:

- Y is the output (educational achievement)
- A is the total factor productivity
- C and N are the inputs (Cognition and Noncognitive skills)
- α and β are the direct effects of inputs
- γ_1 and γ_2 capture curvature (nonlinearities), and γ_{12} captures interaction effects between cognition and noncognitive skills.

6.7.1 Marginal Products (MPs)

$$f_C = \left. \frac{\partial Y}{\partial C} \right|_{N=N_0} = A\alpha C^{\alpha-1} N_0^\beta \frac{\partial}{\partial C} [\exp \{X(C, N_0)\}] \quad (41)$$

$$f_N = \left. \frac{\partial Y}{\partial N} \right|_{C=C_0} = AC_0^\alpha \beta N^{\beta-1} \frac{\partial}{\partial N} [\exp \{X(C_0, N)\}] \quad (42)$$

where $X(C, N)$ is the exponential term in the original function:

$$X(C, N) = \frac{1}{2} \gamma_1 [\ln(C)]^2 + \frac{1}{2} \gamma_2 [\ln(N)]^2 + \gamma_{12} \ln(C) \ln(N) \quad (43)$$

Applying the chain rule to the exponential component and combining terms yields the following expressions for the marginal products, which can be compactly written using the output elasticities derived in Section 7.1.2:

$$f_C = \frac{Y}{C} (\alpha + \gamma_1 \ln(C) + \gamma_{12} \ln(N_0)) = \frac{Y}{C} \cdot OE_C \quad (44)$$

$$f_N = \frac{Y}{N} (\beta + \gamma_2 \ln(N) + \gamma_{12} \ln(C_0)) = \frac{Y}{N} \cdot OE_N \quad (45)$$

6.7.2 Output Elasticities (OEs)

$$\ln Y = \ln A + \alpha \ln C + \beta \ln N + \frac{1}{2} \gamma_1 (\ln C)^2 + \frac{1}{2} \gamma_2 (\ln N)^2 + \gamma_{12} \ln C \ln N \quad (46)$$

To derive output elasticities, we take the partial derivatives of $\ln(Y)$ with respect to $\ln(C)$ and $\ln(N)$:

For C:

$$OE_C = \left. \frac{\partial \ln(Y)}{\partial \ln(C)} \right|_{N=N_0} \quad (47)$$

$$= \frac{\partial}{\partial \ln(C)} \left[\ln A + \alpha \ln C + \beta \ln N_0 + \frac{1}{2} \gamma_1 (\ln C)^2 + \frac{1}{2} \gamma_2 (\ln N_0)^2 + \gamma_{12} \ln C \ln N_0 \right] \quad (48)$$

$$= \alpha + \gamma_1 \ln(C) + \gamma_{12} \ln(N_0) \quad (49)$$

For N:

$$OE_N = \left. \frac{\partial \ln(Y)}{\partial \ln(N)} \right|_{C=C_0} \quad (50)$$

$$= \frac{\partial}{\partial \ln(N)} \left[\ln A + \alpha \ln C_0 + \beta \ln N + \frac{1}{2} \gamma_1 (\ln C_0)^2 + \frac{1}{2} \gamma_2 (\ln N)^2 + \gamma_{12} \ln C_0 \ln N \right] \quad (51)$$

$$= \beta + \gamma_2 \ln(N) + \gamma_{12} \ln(C_0) \quad (52)$$

Therefore, the output elasticities are:

$$OE_C = \left. \frac{\partial \ln(Y)}{\partial \ln(C)} \right|_{N=N_0} = \alpha + \gamma_1 \ln(C) + \gamma_{12} \ln(N_0) \quad (53)$$

$$OE_N = \left. \frac{\partial \ln(Y)}{\partial \ln(N)} \right|_{C=C_0} = \beta + \gamma_2 \ln(N) + \gamma_{12} \ln(C_0) \quad (54)$$

6.7.3 Advantages

The translog model provides a more flexible functional form compared to the Cobb-Douglas models discussed in previous sections, which allows us to capture nonlinear relationships and interactions between cognitive and noncognitive inputs that were not possible in simpler specifications. Some of the key features of these output elasticities are: variable elasticities, input interactions, and nonlinearity. Unlike in the Cobb-Douglas model where elasticities are constant, in the translog model, elasticities vary with the levels of inputs. The cross-term γ_{12} reflects how the marginal effect of one input depends on the level of the other input: $\gamma_{12} \ln(N)$ appears in OE_C , and $\gamma_{12} \ln(C)$ appears in OE_N .

These features make the Translog function useful for modeling non-obvious educational production processes where the impacts of cognitive and noncognitive skills may vary at different levels and interact with each other.

6.7.4 elasticity of substitution

The elasticity of substitution for our translog function is derived from the Marginal Rate of Technical Substitution (MRTS):

$$\sigma = 1 - \frac{\partial \ln(MRTS)}{\partial \ln(C/N)} \quad (55)$$

The MRTS is defined as the ratio of marginal products:

$$MRTS = \frac{\partial Y / \partial C}{\partial Y / \partial N} = \frac{OE_C}{OE_N} \cdot \frac{N}{C} \quad (56)$$

Taking the natural logarithm:

$$\ln(MRTS) = \ln(OE_C) - \ln(OE_N) + \ln(N) - \ln(C) \quad (57)$$

The elasticity of substitution is obtained by differentiating $\ln(MRTS)$ with respect to $\ln(C/N)$, as follows:

a) First, we differentiate $\ln(MRTS)$ with respect to $\ln(C/N)$:

$$\frac{\partial \ln(MRTS)}{\partial \ln(C/N)} = \frac{\partial}{\partial \ln(C/N)} [\ln(OE_C) - \ln(OE_N) + \ln(N) - \ln(C)] \quad (58)$$

b) Using the chain rule and noting that:

$$\frac{\partial \ln(OE_C)}{\partial \ln(C)} = \frac{\gamma_1}{OE_C} \frac{\partial \ln(OE_C)}{\partial \ln(N)} = \frac{\gamma_{12}}{OE_C} \frac{\partial \ln(OE_N)}{\partial \ln(C)} = \frac{\gamma_{12}}{OE_N} \frac{\partial \ln(OE_N)}{\partial \ln(N)} = \frac{\gamma_2}{OE_N} \quad (59)$$

c) Substituting and collecting terms:

$$\frac{\partial \ln(MRTS)}{\partial \ln(C/N)} = \left(\frac{OE_C + OE_N - \gamma_{12} \left(\frac{OE_C}{OE_N} + \frac{OE_N}{OE_C} \right)}{OE_C + OE_N} \right) - 1 \quad (60)$$

Finally, substituting into the original formula:

$$\sigma = 1 - \left[\frac{OE_C + OE_N - \gamma_{12} \left(\frac{OE_C}{OE_N} + \frac{OE_N}{OE_C} \right)}{OE_C + OE_N} - 1 \right] \quad (61)$$

Which simplifies to our final expression:

$$\sigma = \frac{OE_C + OE_N}{OE_C + OE_N - \gamma_{12} \left(\frac{OE_C}{OE_N} + \frac{OE_N}{OE_C} \right)} \quad (62)$$

Where the output elasticities are:

$$OE_C = \alpha + \gamma_1 \ln(C) + \gamma_{12} \ln(N) \quad (63)$$

$$OE_N = \beta + \gamma_2 \ln(N) + \gamma_{12} \ln(C) \quad (64)$$

This formulation captures three key features:

- Total productivity: The numerator ($OE_C + OE_N$) measures combined skill contribution
- Skill interaction: The γ_{12} term captures complementarity effects
- Skill balance: The ratios $\frac{OE_C}{OE_N}$ and $\frac{OE_N}{OE_C}$ reflect relative skill intensity

The substitutability condition is:

$$\gamma_{12} \left(\frac{OE_C}{OE_N} + \frac{OE_N}{OE_C} \right) > 0 \quad (65)$$

When this holds, cognitive and noncognitive skills exhibit diminishing substitutability in educational achievement production.

6.7.5 Marginal Rate of Technical Substitution (MRTS)

1. By definition, $MRTS_{CN} = \frac{f_C}{f_N}$
2. We know that $f_C = OE_C \cdot \frac{Y}{C}$ and $f_N = OE_N \cdot \frac{Y}{N}$
3. Therefore:

$$MRTS_{CN} = \frac{f_C}{f_N} = \frac{OE_C \cdot \frac{Y}{C}}{OE_N \cdot \frac{Y}{N}} = \frac{OE_C}{OE_N} \cdot \frac{N}{C} \quad (66)$$

4. Substituting the expressions for OE_C and OE_N :

$$MRTS_{CN} = \frac{\alpha + \gamma_1 \ln(C) + \gamma_{12} \ln(N)}{\beta + \gamma_2 \ln(N) + \gamma_{12} \ln(C)} \cdot \frac{N}{C} \quad (67)$$

6.8 General Form: CES with Two Inputs

$$Y = A [\alpha C^\rho + (1 - \alpha) N^\rho]^{\frac{1}{\rho}} \quad (68)$$

Where:

Y : Total output/Grade function/Academic achievement

A : Total factor productivity or scaling factor

α : Share parameter for cognitive input

ρ : Substitution parameter, where $\rho = \frac{\sigma - 1}{\sigma}$

σ : Elasticity of substitution

C : Input representing cognition

N : Input representing noncognitive measure

6.8.1 Elasticity of Substitution

The elasticity of substitution (σ) measures how easily cognitive and noncognitive inputs can be substituted for each other. The relationship between σ and ρ governs how easily the two inputs can substitute for each other:

- When $\sigma > 1$ (or $-1 < \rho < \infty$), cognitive and noncognitive inputs are substitutes.
- When $\sigma < 1$ (or $\rho < -1$), cognitive and noncognitive inputs are complements.
- As σ approaches infinity (or ρ approaches 1), the inputs become perfect substitutes.
- As σ approaches 0 (or ρ approaches $-\infty$), the inputs become perfect complements.
- When $\sigma = 1$ (or equivalently, $\rho = 0$), the CES collapses to the Cobb-Douglas form.

In the context of educational production, these links clarify how cognitive and noncognitive skills interact to shape academic outcomes. For example, when $\sigma > 1$, a deficiency in one skill type can be more easily compensated by the other.

6.8.2 Marginal Products (MPs)

$$f_C = \frac{\partial Y}{\partial C} = A\alpha C^{\rho-1} [\alpha C^\rho + (1 - \alpha)N^\rho]^{\frac{1}{\rho}-1} \quad (69)$$

$$f_N = \frac{\partial Y}{\partial N} = A(1 - \alpha)N^{\rho-1} [\alpha C^\rho + (1 - \alpha)N^\rho]^{\frac{1}{\rho}-1} \quad (70)$$

6.8.3 Output elasticities (OEs)

$$OE_C = \frac{\partial \ln(Y)}{\partial \ln(C)} = \frac{\alpha C^\rho}{\alpha C^\rho + (1 - \alpha)N^\rho} \quad (71)$$

$$OE_N = \frac{\partial \ln(Y)}{\partial \ln(N)} = \frac{(1 - \alpha)N^\rho}{\alpha C^\rho + (1 - \alpha)N^\rho} \quad (72)$$

The sum of output elasticities would still be 1, indicating constant returns to scale:

$$OE_C + OE_N = \frac{\alpha C^\rho + (1 - \alpha)N^\rho}{\alpha C^\rho + (1 - \alpha)N^\rho} = 1 \quad (73)$$

This holds regardless of the values of C , N , α , or ρ , confirming that the CES function exhibits constant returns to scale by construction.

6.8.4 Returns to scale

The interpretation of constant returns to scale remains the same as in the three-input case: a proportional increase in both cognitive and noncognitive inputs leads to an equivalent proportional increase in the educational output.

The degree of returns to scale is determined by the sum of all output elasticities. We can call this sum the scale elasticity (SE):

$$SE = OE_C + OE_N \quad (74)$$

Then:

- a) If $SE > 1$ = Increasing returns to scale;
- b) If $SE < 1$ = Decreasing returns to scale;
- c) If $SE = 1$ = Constant returns to scale.

For:

$$Y = A [\alpha C^\rho + (1 - \alpha)N^\rho]^{\frac{1}{\rho}} \quad (75)$$

The sum of the output elasticities is always 1, regardless of the parameter values:

$$OE_C + OE_N = \frac{\alpha C^\rho + (1 - \alpha)N^\rho}{\alpha C^\rho + (1 - \alpha)N^\rho} = 1 \quad (76)$$

the 2-input CES function exhibits constant returns to scale by construction. This is a property of the CES function with the exponent $\frac{1}{\rho}$ outside the brackets. In the context of cognition and noncognition as inputs in an educational production function, it means that if both inputs are scaled by a constant factor k , then output Y increases proportionally by the same factor. More specifically:

1. Proportional increase in inputs:

$$C \rightarrow kC, \quad N \rightarrow kN \quad (77)$$

2. Resulting increase in output:

$$Y(kC, kN) = kY(C, N) \quad (78)$$

In practical terms for education, this means a proportional improvement in cognitive and noncognitive skills leads to an equivalent proportional improvement in educational outcomes.

If we were to allow for different returns to scale, we could modify the CES function to:

$$Y = A [\alpha C^\rho + (1 - \alpha)N^\rho]^{\frac{\nu}{\rho}} \quad (79)$$

Where ν is a new parameter that determines the overall returns to scale:

- a) If $\nu > 1$ = Increasing returns to scale;
- b) If $\nu < 1$ = Decreasing returns to scale;
- c) If $\nu = 1$ = Constant returns to scale (current case).

6.8.5 Marginal Rate of Technical Substitution (MRTS)

For the two-input CES production function:

$$Q = A (\alpha C^\rho + (1 - \alpha)N^\rho)^{\frac{1}{\rho}} \quad (80)$$

The marginal products are:

$$f_C = \frac{\partial Q}{\partial C} = A\alpha (\alpha C^\rho + (1 - \alpha)N^\rho)^{\frac{1}{\rho}-1} C^{\rho-1} \quad (81)$$

$$f_N = \frac{\partial Q}{\partial N} = A(1 - \alpha) (\alpha C^\rho + (1 - \alpha)N^\rho)^{\frac{1}{\rho}-1} N^{\rho-1} \quad (82)$$

$$MRTS_{CN} = \frac{f_C}{f_N} = \frac{\alpha C^{\rho-1}}{(1 - \alpha)N^{\rho-1}} = \frac{\alpha}{1 - \alpha} \left(\frac{C}{N} \right)^{\rho-1} \quad (83)$$

In education terms, this shows how much noncognitive skill is needed to replace one unit of cognition (or vice versa), depending on their relative levels and substitutability.

6.8.6 Isoquants

Isoquants for the two-input CES production function represent combinations of C and N that produce the same level of output Y . For the two-input case, we can represent isoquants as follows:

1. Equation form:

For a given output level Y_0 , the isoquant is represented by:

$$Y_0 = A [\alpha C^\rho + (1 - \alpha)N^\rho]^{1/\rho} \quad (84)$$

This can be rearranged to express N in terms of C :

$$N = \left[\frac{(Y_0^\rho/A^\rho) - \alpha C^\rho}{1 - \alpha} \right]^{1/\rho} \quad (85)$$

2. Graphical representation:

In the two-dimensional space of C and N , each isoquant is a curve representing all combinations of cognitive and noncognitive inputs that produce the same level of output Y_0 .

The shape of the isoquants reflects the substitutability between cognitive and noncognitive inputs:

- As ρ approaches 1 (or σ approaches infinity), the isoquants become more linear, indicating that C and N are close to perfect substitutes.
- As ρ approaches negative infinity (or σ approaches 0), the isoquants approach right angles, indicating that C and N are close to perfect complements.
- When $\rho = 0$ (or $\sigma = 1$), the isoquants take on the familiar Cobb-Douglas shape.

These isoquant properties help visualize how cognitive and noncognitive skills can be substituted in different proportions to achieve the same academic outcome.

6.9 Limitations

While the two-input CES model further enhances our understanding of the connection between cognitive and noncognitive skills in educational production, I have to note it has some limitations. Regarding noncognitive skill selection, the model treats noncognition as a single aggregate input, which may oversimplify its multidimensional nature. The assumption of constant elasticity of substitution may not hold uniformly across different levels of input use. Other factors that influence educational outcomes, such as family background or school quality, are not explicitly included in this model (but are in the regressions).

6.10 General Form: CES with Three Inputs

$$Y = A[\alpha C^\rho + (1 - \alpha)(\beta N_E^\rho + (1 - \beta)N_I^\rho)]^{\frac{1}{\rho}} \quad (86)$$

Where:

Y : Total output/Grade function/Academic achievement

A : Total factor productivity or scaling factor

α : Share parameter for cognitive input

β : Share parameter for noncognitive inputs

ρ : Substitution parameter, where $\rho = \frac{\sigma - 1}{\sigma}$ (i.e., $\sigma = \frac{1}{1 - \rho}$)

σ : Elasticity of substitution

C : Input representing cognition

N_E, N_I : Inputs representing noncognitive measures

This is a three-input Constant Elasticity of Substitution (CES) production function, which generalizes the two-input case by allowing differentiated treatment of multiple noncognitive dimensions. The CES function is more flexible than the Cobb-Douglas form, allowing for varying degrees of substitutability between inputs. The elasticity of substitution (σ) between any pair of inputs is constant and determined by the parameter ρ .

In this model, C represents a measure of cognition (in this case, the principal component as a composite of three cognitive measures). N_E and N_I represent noncognitive measures, which I call External Control and Internal Control, respectively. In relation to the scales used (TIPI and SDQ), Internal Control proxies Focused Behaviour (SDQ) and Conscientiousness (TIPI), while External Control captures Emotional Resilience (SDQ) and Emotional Stability (TIPI). These four variables appear to be the most significant based on my analysis.

The share parameters α and β determine the relative importance of the inputs in the production function. However, unlike in a Cobb-Douglas function, they do not directly determine output elasticities, which vary with the levels of input usage in the CES model.

6.10.1 Elasticity of Substitution

The elasticity of substitution (σ) in the three-input CES model measures the ease of substitution between any pair of inputs while holding the third input constant. The relationship between σ and ρ is:

- When $\sigma > 1$ (or $-1 < \rho < \infty$), any pair of inputs are substitutes.

- When $\sigma < 1$ (or $\rho < -1$), any pair of inputs are complements.
- As σ approaches infinity (or ρ approaches 1), the inputs become perfect substitutes.
- As σ approaches 0 (or ρ approaches $-\infty$), the inputs become perfect complements.
- When $\sigma = 1$ (or $\rho = 0$), the CES function reduces to the Cobb-Douglas form.

In the context of educational production with cognitive (C), external noncognitive (N_E), and internal noncognitive (N_I) inputs, these relationships indicate how these different skills interact in producing educational outcomes. For example:

- If $\sigma > 1$, a deficiency in one type of skill (e.g., cognitive) can be more easily compensated by either of the other skills.
- If $\sigma < 1$, it implies that all three types of skills are complementary, and a balanced development of all skills is important for educational outcomes.

Although σ reflects the overall substitutability, the actual trade-offs between input pairs (e.g., C vs. N_E) depend on both the parameter values and the relative levels of the inputs. In the proposed model, I assume a constant elasticity of substitution between all input pairs, which is a simplification of potentially nonlinear relationships in real-life educational production.

While this assumption improves tractability and interpretability, it is worth noting that more flexible nested CES forms allow for different elasticities of substitution between input pairs. For example, the elasticity between cognitive and noncognitive inputs could differ from that between the two noncognitive dimensions. Such models offer richer behavioural insights but also involve substantially greater complexity and identification challenges. For this analysis, I maintain a constant σ to preserve parsimony and enable clearer comparisons across specifications.

6.10.2 Marginal products (MPs)

$$f_C = \left. \frac{\partial Y}{\partial C} \right|_{N_E=N_{E0}, N_I=N_{I0}} = A\alpha C^{\rho-1} [\alpha C^\rho + (1-\alpha)(\beta N_{E0}^\rho + (1-\beta)N_{I0}^\rho)]^{\frac{1}{\rho}-1} \quad (87)$$

$$f_{N_E} = \left. \frac{\partial Y}{\partial N_E} \right|_{C=C_0, N_I=N_{I0}} = A(1-\alpha)\beta N_E^{\rho-1} [\alpha C_0^\rho + (1-\alpha)(\beta N_E^\rho + (1-\beta)N_{I0}^\rho)]^{\frac{1}{\rho}-1} \quad (88)$$

$$f_{N_I} = \left. \frac{\partial Y}{\partial N_I} \right|_{C=C_0, N_E=N_{E0}} = A(1-\alpha)(1-\beta)N_I^{\rho-1} [\alpha C_0^\rho + (1-\alpha)(\beta N_{E0}^\rho + (1-\beta)N_I^\rho)]^{\frac{1}{\rho}-1} \quad (89)$$

6.10.3 Output elasticities (OEs)

Given:

$$Y = A [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}} \quad (90)$$

When we take the log on both sides:

$$\ln(Y) = \ln(A) + \frac{1}{\rho} \ln [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)] \quad (91)$$

This yields the following output elasticities for each input:

$$OE_C = \left. \frac{\partial \ln(Y)}{\partial \ln(C)} \right|_{N_E=N_{E0}, N_I=N_{I0}} = \frac{\alpha C^\rho}{\alpha C^\rho + (1 - \alpha) (\beta N_{E0}^\rho + (1 - \beta) N_{I0}^\rho)} \quad (92)$$

$$OE_{N_E} = \left. \frac{\partial \ln(Y)}{\partial \ln(N_E)} \right|_{C=C_0, N_I=N_{I0}} = \frac{(1 - \alpha) \beta N_E^\rho}{\alpha C_0^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_{I0}^\rho)} \quad (93)$$

$$OE_{N_I} = \left. \frac{\partial \ln(Y)}{\partial \ln(N_I)} \right|_{C=C_0, N_E=N_{E0}} = \frac{(1 - \alpha) (1 - \beta) N_I^\rho}{\alpha C_0^\rho + (1 - \alpha) (\beta N_{E0}^\rho + (1 - \beta) N_I^\rho)} \quad (94)$$

Step-by-step derivation:

1) First we take the natural logarithm:

$$\ln(Y) = \ln(A) + \frac{1}{\rho} \ln [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)] \quad (95)$$

2) Then we derive the output elasticities one by one:

For Cognitive input (C):

$$OE_C = \frac{\partial \ln(Y)}{\partial \ln(C)} = \frac{\alpha C^\rho}{\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)} \quad (96)$$

For External Noncognitive input (N_E):

$$OE_{N_E} = \frac{\partial \ln(Y)}{\partial \ln(N_E)} = \frac{(1 - \alpha) \beta N_E^\rho}{\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)} \quad (97)$$

For Internal Noncognitive input (N_I):

$$OE_{N_I} = \frac{\partial \ln(Y)}{\partial \ln(N_I)} = \frac{(1 - \alpha) (1 - \beta) N_I^\rho}{\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)} \quad (98)$$

6.10.4 Returns to scale

The degree of returns to scale is determined by the sum of all output elasticities. We can call this sum the scale elasticity (SE):

$$SE = OE_C + OE_{N_E} + OE_{N_I} \quad (99)$$

Then:

- a) If $SE > 1$ = Increasing returns to scale;
- b) If $SE < 1$ = Decreasing returns to scale;
- c) If $SE = 1$ = Constant returns to scale.

For:

$$Y = A [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}} \quad (100)$$

The sum of the output elasticities is always 1, regardless of the parameter values:

$$OE_C + OE_{N_E} + OE_{N_I} = \frac{\alpha C^\rho + (1 - \alpha) \beta N_E^\rho + (1 - \alpha) (1 - \beta) N_I^\rho}{\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)} = 1 \quad (101)$$

This 3-input CES function exhibits constant returns to scale by construction. This is a property of the CES function with the exponent $\frac{1}{\rho}$ outside the brackets. In the context of cognition and noncognition as inputs in an educational production function, it means that if we increase all three inputs by a factor k then output Y increases proportionally by the same factor. More specifically:

1. Proportional increase in inputs:

$$C \rightarrow kC, \quad N_E \rightarrow kN_E, \quad N_I \rightarrow kN_I \quad (102)$$

2. Resulting increase in output:

$$Y(kC, kN_E, kN_I) = kY(C, N_E, N_I) \quad (103)$$

In practical terms for education, this means a proportional improvement in cognitive and noncognitive skills leads to an equivalent proportional improvement in educational outcomes. For example, if we could somehow double ($k = 2$) a student's cognitive ability (C) and both types of noncognitive abilities (N_E and N_I) simultaneously, we would expect their educational output (Y , measured by test scores as a proxy for overall academic performance) to also double. This implies:

- a) No diminishing returns when scaling up all inputs equally;
- b) No extra benefits (increasing returns) when scaling up all inputs equally.

We need to keep in mind that this is a simplification of a sophisticated reality. In practice, the links between cognitive abilities, noncognitive skills, and educational outcomes

is more nuanced and also most likely nonlinear, as we have seen in previous chapters.

If we were to allow for different returns to scale, we could generalize the CES form to allow variable returns to scale:

$$Y = A [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{\nu}{\rho}} \quad (104)$$

Where ν is a new parameter that determines the overall returns to scale:

- a) If $\nu > 1$ = Increasing returns to scale;
- b) If $\nu < 1$ = Decreasing returns to scale;
- c) If $\nu = 1$ = Constant returns to scale (current case).

6.10.5 Marginal Rate of Technical Substitution for Three-Input CES

For the three-input CES production function:

$$Q = A [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}} \quad (105)$$

The marginal products are:

$$f_C = \frac{\partial Q}{\partial C} = A \alpha C^{\rho-1} [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}-1} \quad (106)$$

$$f_{N_E} = \frac{\partial Q}{\partial N_E} = A (1 - \alpha) \beta N_E^{\rho-1} [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}-1} \quad (107)$$

$$f_{N_I} = \frac{\partial Q}{\partial N_I} = A (1 - \alpha) (1 - \beta) N_I^{\rho-1} [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}-1} \quad (108)$$

The MRTS can be calculated for each pair of inputs. Holding the third input constant, these expressions quantify how much of one input is required to offset a marginal decrease in another while maintaining the same output level:

1. Between cognitive (C) and external noncognitive (N_E) inputs:

$$MRTS_{C,N_E} = \frac{f_C}{f_{N_E}} = \frac{\alpha}{\beta(1 - \alpha)} \left(\frac{N_E}{C} \right)^{1-\rho} \quad (109)$$

2. Between cognitive (C) and internal noncognitive (N_I) inputs:

$$MRTS_{C,N_I} = \frac{f_C}{f_{N_I}} = \frac{\alpha}{(1 - \beta)(1 - \alpha)} \left(\frac{N_I}{C} \right)^{1-\rho} \quad (110)$$

3. Between external noncognitive (N_E) and internal noncognitive (N_I) inputs:

$$MRTS_{N_E, N_I} = \frac{f_{N_E}}{f_{N_I}} = \frac{\beta}{1 - \beta} \left(\frac{N_I}{N_E} \right)^{1-\rho} \quad (111)$$

These MRTS formulas demonstrate how the substitutability between each pair of inputs changes with their relative quantities and the elasticity of substitution parameter ρ . We can analyze the trade-offs between any two of the three inputs while holding the third constant.

For example, $MRTS_{C, N_E}$ shows how much external noncognitive input (N_E) is needed to compensate for a small decrease in cognitive input (C) while maintaining the same output level and holding internal noncognitive input (N_I) constant. The substitutability is governed by the share parameters (α, β), the substitution parameter (ρ), and the current relative input levels.

6.10.6 Isoquants

Isoquants represent all input combinations that yield the same level of output (Y). For the three-input CES function, these curves - or surfaces - illustrate the substitution possibilities among C , N_E , and N_I . Due to the three-dimensional nature of the input space, we can represent isoquants in a few ways:

1. Two-dimensional representation:

Fixing C at a level C_0 , we can represent the isoquant for output level Y_0 as:

$$N_I = \left[\frac{(Y_0^\rho / A^\rho - \alpha C_0^\rho)}{(1 - \alpha)} - \beta N_E^\rho \right]^{1/\rho} / (1 - \beta)^{1/\rho} \quad (112)$$

This expression defines the isoquant curve in the (N_E, N_I) plane for a fixed cognitive level C_0 and output level Y_0 .

2. Three-dimensional representation:

The full isoquant surface for output level Y_0 is given by:

$$Y_0 = A [\alpha C^\rho + (1 - \alpha)(\beta N_E^\rho + (1 - \beta)N_I^\rho)]^{1/\rho} \quad (113)$$

This surface in (C, N_E, N_I) space represents all combinations of inputs producing output Y_0 .

The shape of the isoquants reflects the substitutability between inputs. As ρ approaches 1 (perfect substitutes), the isoquants become more linear. As ρ approaches negative infinity (perfect complements), the isoquants approach right angles.

6.11 Comparison Between Two-Inputs and Three-Inputs CES Models

The choice between the two-input and three-input CES models involves several trade-offs, and was temporarily set aside due to empirical (primarily computational) limitations:

Simplicity vs. complexity: The two-input model offers greater simplicity and ease of interpretation, making it more suitable for theoretical analysis and empirical estimation. However, the three-input model provides a more nuanced representation of noncognitive skills - distinguishing between internal and external control - and enables the researcher to select the most relevant inputs. If better measures existed on a unified scale, cross-model comparisons would be more straightforward. Currently, differences between the TIPI and SDQ scales constrain the use of principal component analysis.

Parsimony vs. granularity: The two-input model is more parsimonious, requiring fewer parameters to estimate (one for cognitive and one for noncognitive input). This is advantageous when working with limited or noisy data, especially for noncognitive variables. The three-input model, although more complex, offers greater granularity in modeling educational production processes. Both have merits and limitations, and the choice should depend on the research context and goals.

Generalizability vs. specificity: The two-input model may be more generalizable across settings where noncognitive dimensions are not easily separable. In contrast, the three-input model is more suitable for contexts where distinct aspects of noncognition-such as internal and external control-are meaningfully identified and measured.

Ultimately, the choice depends on the research question, data quality, and patience for experimenting with different optimization routines when setting starting values for estimation.

6.11.1 Nested CES function

While the previous analysis focuses on the standard CES function, it is worth noting the possibility of using a nested CES function, particularly for the three-input case. The first specification was:

$$Y = A [\alpha C^\rho + (1 - \alpha) (\beta N_E^\rho + (1 - \beta) N_I^\rho)]^{\frac{1}{\rho}} \quad (114)$$

While a nested CES function could look like:

$$Y = A \left[\alpha C^\rho + (1 - \alpha) (\beta N_E^\gamma + (1 - \beta) N_I^\gamma)^{\rho/\gamma} \right]^{\frac{1}{\rho}} \quad (115)$$

This nested specification allows for different elasticities of substitution between cognitive and noncognitive inputs (determined by ρ) and between the two types of noncognitive

inputs (determined by γ).

This additional parameter, γ , allows for different elasticities of substitution between inputs:

- ρ determines the elasticity of substitution between cognitive skills (C) and the composite of noncognitive skills (N_E and N_I).
- γ determines the elasticity of substitution between the two types of noncognitive skills (N_E and N_I).
- When $\gamma > \rho$, the two noncognitive inputs (N_E and N_I) are more substitutable with each other than either is with the cognitive input (C).
- When $\gamma < \rho$, the noncognitive inputs are less substitutable with each other than with the cognitive input.
- When $\gamma = \rho$, the nested CES function reduces to the standard three-input CES function.

This nested structure is more interesting at first because it allows for detailed modeling of the links between different types of skills. For example, it can capture scenarios where external and internal noncognitive skills might be more easily substituted for each other than either can be for cognitive skills.

In educational terms, a high γ relative to ρ might suggest that deficiencies in one type of noncognitive skill (e.g., external control) can be more easily compensated by strengths in the other noncognitive skill (e.g., internal control) than by cognitive abilities.

However, while this alternative model could capture more nuanced links between the inputs, it would be even more computationally complex to estimate, and the researcher would have to justify the choice of variables.

6.11.2 Economic interpretation of parameters

The parameters in the previous CES models have important economic interpretations in the context of educational production:

- α (and β in the three-input case) govern the relative weighting of inputs in the production process. A higher α indicates that cognitive skills have a greater weight relative to noncognitive skills in determining educational outcomes. These parameters reflect the underlying production technology.
- ρ , which maps one-to-one to the elasticity of substitution σ via $\sigma = 1/(1 - \rho)$, governs the degree of substitutability between inputs. In educational terms, it reflects how easily a deficiency in one type of skill can be compensated by strength

in another. A higher σ reflects greater flexibility in combining different skills to achieve educational outcomes.

- A represents total factor productivity, which in an educational context might reflect the overall effectiveness of the educational system or other factors that affect all students equally.

6.11.3 Interpretive implications

The CES framework provides an interpretation guide for how skill margins may combine. Starting with the elasticity of substitution (σ), if it is high, model-implied trade-offs suggest that deficiencies in one area can be more easily offset by strengths in another. If σ is low, the model implies that balanced skill profiles are more strongly associated with achievement.

The relative magnitudes of α and β provide a compact summary of which input carries a steeper marginal association in the fitted model. These are descriptive model implications, not treatment-effect estimates.

The returns-to-scale properties indicate whether proportional increases in both inputs are associated with proportional, less than proportional, or greater than proportional changes in predicted achievement.

Prior literature indicates that cognitive and noncognitive development differ in timing and malleability, with important interactions between endowments and environments (Kautz et al., 2014; Knudsen et al., 2006). In this chapter, these references provide interpretation context rather than causal evidence on programme effectiveness.

The early childhood (0-5 years) period is often highlighted as a sensitive period for cognitive development, although the brain retains plasticity throughout life (Knudsen et al., 2006). Evidence from intervention studies such as Perry Preschool and Abecedarian is relevant context for interpretation, but those designs are separate from the observational framework used here (Campbell et al., 2012; Schweinhart et al., 2005).

Related intervention evidence also suggests that some noncognitive margins can respond at later stages, including during school years (Raver et al., 2011). Again, this informs external interpretation rather than identification within this chapter.

The work of James Heckman and colleagues documents that early interventions can have persistent effects on both cognitive and noncognitive skills, and that noncognitive margins may remain more malleable later in life (Heckman & Kautz, 2012; Kautz et al., 2014).

It is also important to note that developmental outcomes reflect interactions between endowments and environments (Fox et al., 2010). Evidence from other designs, including nurse home visiting programmes, provides complementary context on possible mechanisms (Olds et al., 2004).

Taken together, these references offer a plausible interpretation framework for the empirical patterns in this chapter, but they do not turn the present estimates into causal policy effects. The chapter’s contribution remains descriptive: mapping how observed cognitive and noncognitive skill bundles co-vary with achievement by subject and gender in the GUI cohort.

6.12 Empirical considerations

The CES specification offers a theoretically appealing framework, but its empirical implementation encountered several difficulties. The CES function is nonlinear in its parameters, and nonlinear least squares estimation exhibited numerical instability in several configurations - including boundary solutions when the substitution parameter approached zero and non-convergence when the ratio of cognitive to noncognitive inputs varied little across observations. These identification challenges motivated the translog specification used in the main analysis, which nests the CES as a special case while permitting standard linear-in-parameters estimation.

Measurement error remains a substantive concern. Both cognitive and noncognitive inputs are measured with error: cognitive ability through standardised tests that capture only specific facets of general ability, and noncognitive traits through primary caregiver ratings that reflect home-observed behaviour rather than classroom conduct. Classical measurement error in either input would attenuate the estimated complementarity, biasing the elasticity of substitution toward unity. The finding that ES is significantly below one under SDQ specifications is therefore conservative with respect to this source of bias.

Finally, all production-function specifications impose functional-form restrictions on the mapping from latent skills to observed achievement. Educational achievement is not directly observable - it is approximated through bounded, discrete test-based proxies. The robustness checks in the main text (z-score outcome analogues, top-decile trimming, alternative evaluation points, and Leaving Certificate outcome extensions) are designed to ensure that the complementarity conclusion is not driven by the specific functional form or outcome scale.

6.13 Selection and attrition diagnostics

Table 20 compares baseline characteristics of retained versus lost observations at two selection margins. The retained analytical sample is positively selected on prior achievement and socioeconomic status, but effect sizes are moderate (Cohen’s $d < 0.35$ for all variables). Table 21 reports inverse-probability-weighted OLS interaction coefficients: the primary SDQ Focused Behaviour \times Cognition interaction is robust to selection reweighting, with coefficient changes of 3–11%.

Table 20: Baseline balance: retained vs. lost observations at two selection margins. Panel A compares Wave 1 characteristics of the full GUI cohort ($N = 8,568$) by retention in the analytical sample ($N = 5,631$). Panel B compares the analytical sample by retention in the restricted (complete-controls) sample ($N \approx 3,786$). Cohen’s d is the standardised mean difference (positive = retained group higher).

Variable	Retained		Lost		d	p
	Mean	SD	Mean	SD		
<i>Panel A: Full cohort → analytical sample</i>						
Reading (% correct, W1)	73.10	(19.86)	65.88	(21.79)	+0.351	0.000
Maths (% correct, W1)	58.82	(20.44)	51.76	(21.28)	+0.341	0.000
SDQ Emotional symptoms (W1)	1.96	(1.94)	2.11	(2.03)	-0.072	0.002
SDQ Conduct problems (W1)	1.18	(1.38)	1.39	(1.51)	-0.144	0.000
SDQ Hyperactivity (W1)	2.87	(2.37)	3.21	(2.51)	-0.143	0.000
SDQ Peer problems (W1)	1.09	(1.41)	1.24	(1.47)	-0.101	0.000
SDQ Total difficulties (W1)	7.11	(4.86)	7.94	(5.32)	-0.167	0.000
Income quintile (equiv., W1)	3.43	(1.34)	3.11	(1.43)	+0.230	0.000
<i>Panel B: Analytical → restricted (complete controls) sample</i>						
Reading (% correct, W1)	74.46	(19.23)	70.30	(20.84)	+0.210	0.000
Maths (% correct, W1)	60.17	(20.08)	56.04	(20.88)	+0.203	0.000
SDQ Emotional res. (W1)	8.12	(1.89)	7.86	(2.04)	+0.136	0.000
SDQ Good conduct (W1)	8.86	(1.35)	8.73	(1.45)	+0.098	0.001
SDQ Focused behav. (W1)	7.19	(2.35)	7.02	(2.41)	+0.070	0.014
SDQ Pos. peers (W1)	8.98	(1.36)	8.75	(1.50)	+0.159	0.000
Maths points (JC outcome)	9.75	(1.68)	9.30	(1.81)	+0.259	0.000
English points (JC outcome)	10.25	(1.26)	9.94	(1.45)	+0.232	0.000
SDQ Focused behav. (W2)	7.66	(2.21)	7.38	(2.36)	+0.124	0.000
TIPI Conscient. (W2)	4.36	(2.06)	4.26	(2.10)	+0.047	0.097
PCG Education (W2)	4.02	(1.23)	3.84	(1.26)	+0.144	0.000

Note: d is Cohen’s standardised mean difference (positive = retained group higher). p -values from Welch’s two-sample t -test. Panel A uses Wave 1 (age 9) variables available for the full cohort; Panel B uses Wave 1 and Wave 2 variables available within the analytical sample.

Table 21: IPW sensitivity: OLS interaction coefficients with and without inverse-probability weighting for selection into the restricted (complete-controls) sample. Weights are estimated from a logit of sample retention on Wave 1 baseline characteristics, trimmed at the 1st/99th percentiles.

Outcome	Interaction	Unweighted		IPW-weighted		Change (%)
		Coef	SE	Coef	SE	
Maths	emot \times Cog	-0.0430*	(0.0216)	-0.0399+	(0.0224)	+7.2%
Maths	cond \times Cog	0.0172	(0.0240)	0.0186	(0.0249)	+8.2%
Maths	hyper \times Cog	-0.0828***	(0.0231)	-0.0739**	(0.0246)	+10.8%
Maths	peer \times Cog	-0.0199	(0.0192)	-0.0234	(0.0202)	-17.5%
English	emot \times Cog	-0.0263	(0.0208)	-0.0240	(0.0218)	+8.9%
English	cond \times Cog	0.0324	(0.0240)	0.0322	(0.0252)	-0.5%
English	hyper \times Cog	-0.0979***	(0.0232)	-0.1004***	(0.0250)	-2.6%
English	peer \times Cog	-0.0379+	(0.0198)	-0.0389+	(0.0207)	-2.8%

Note: IPW weights are the inverse of $\Pr(\text{in restricted sample} \mid \text{Wave 1 baselines})$, estimated by logit on the full analytical sample ($N = 5,631$) and applied to the restricted sample ($N \approx 3,786$). Weights trimmed at 1st/99th percentiles and normalised to sum to N . HC1 robust SEs. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

6.14 Additional robustness checks

This appendix collects supplementary robustness tables referenced in the main text, organised by the concern each addresses. Table 22 checks whether ES estimates change when the sample is restricted to observations with complete controls. Tables 23 and 24 test sensitivity to outcome scaling (z-scored vs. log) and ceiling effects (trimming the top decile). Tables 25 and 26 assess whether the translog is well behaved across the input space. Tables 27–28 gauge how much selection on unobservables would be needed to eliminate the key interaction effects. Tables 29 and 30 simulate school-level clustering to test whether significance survives without access to school identifiers.

Table 22: Robustness: Full Sample vs Restricted (Complete Controls) Sample - Translog Elasticity of Substitution. The full sample includes all observations with non-missing outcome, cognitive, and noncognitive variables ($N = 5,631$). The restricted sample additionally requires non-missing SES and school controls ($N = 3,786$).

Outcome	Noncog	Full Sample			Restricted Sample		
		ES	95% CI	N	ES	95% CI	N
Maths	TIPI	0.569	[0.354, 1.701]	5,631	0.529	[0.243, 2.692]	3,786
Maths	SDQ	0.482	[0.378, 0.673]	5,631	0.517	[0.370, 0.844]	3,786
English	TIPI	0.514	[0.226, 1.914]	5,631	0.470	[-2.042, 3.264]	3,786
English	SDQ	0.464	[0.349, 0.690]	5,631	0.425	[0.325, 0.615]	3,786

Note: Translog shift = 0.5. Bootstrap 95% percentile CIs from 1,000 replications.

Table 23: Robustness: second-order polynomial with z-scored outcomes vs. log outcomes. The cross-term coefficient ($\hat{\beta}_{C \times N}$) captures the interaction between cognitive and noncognitive inputs. A negative cross-term in the z-score specification is the levels-based analogue of $ES < 1$ in the translog: both indicate that the inputs are complements rather than substitutes.

Outcome	Noncog	Subgroup	$\hat{\beta}_{C \times N}$ (z-score)	95% CI	$\hat{\beta}_{C \times N}$ (log)	Sign match
Maths	SDQ (F.B.)	Full sample	-0.0362***	[-0.058, -0.015]	-0.0105***	✓
Maths	SDQ (F.B.)	Boys	-0.0317*	[-0.062, -0.001]	-0.0096**	✓
Maths	SDQ (F.B.)	Girls	-0.0377*	[-0.069, -0.006]	-0.0110**	✓
Maths	TIPI (C.)	Full sample	-0.0219*	[-0.041, -0.004]	-0.0054*	✓
Maths	TIPI (C.)	Boys	-0.0389**	[-0.066, -0.009]	-0.0097**	✓
Maths	TIPI (C.)	Girls	-0.0056	[-0.034, 0.023]	-0.0017	✓
English	SDQ (F.B.)	Full sample	-0.0559***	[-0.082, -0.030]	-0.0096***	✓
English	SDQ (F.B.)	Boys	-0.0477*	[-0.089, -0.008]	-0.0087**	✓
English	SDQ (F.B.)	Girls	-0.0664***	[-0.106, -0.027]	-0.0105***	✓
English	TIPI (C.)	Full sample	-0.0177	[-0.042, 0.007]	-0.0031	✓
English	TIPI (C.)	Boys	-0.0267	[-0.065, 0.008]	-0.0046	✓
English	TIPI (C.)	Girls	-0.0035	[-0.037, 0.032]	-0.0008	✓

Note: F.B. = Focused Behaviour; C. = Conscientiousness. Second-order polynomial: $z(Y) = \alpha + \beta_1 C + \beta_2 N + \beta_3 C^2 + \beta_4 N^2 + \beta_5 C \times N$, where C = cognition (PC, standardised) and N = noncognitive measure (standardised). Z-score specification uses the standardised outcome (mean 0, SD 1); log specification uses $\ln(Y)$. 95% CIs are percentile bootstrap (1,000 replications). HC1 robust SEs for point estimates. “Sign match” indicates whether the cross-term has the same sign in both specifications. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 24: Ceiling-effect robustness for translog ES: baseline sample versus top-decile-trimmed outcomes. The trimmed specification excludes observations at or above the 90th percentile of the outcome distribution (12 points in both subjects).

Outcome	Noncog input	Sample	N	ES [95% CI]	$p(ES = 1)$
Maths	SDQ (F.B.)	Baseline	5,631	0.482 [0.374, 0.660]	0.000
Maths	SDQ (F.B.)	Trim top decile	5,022	0.478 [0.369, 0.629]	0.000
Maths	TIPI (C.)	Baseline	5,631	0.569 [0.337, 1.695]	0.178
Maths	TIPI (C.)	Trim top decile	5,022	0.571 [0.306, 3.276]	0.230
English	SDQ (F.B.)	Baseline	5,631	0.464 [0.347, 0.656]	0.000
English	SDQ (F.B.)	Trim top decile	4,899	0.414 [0.307, 0.580]	0.000
English	TIPI (C.)	Baseline	5,631	0.514 [0.255, 2.270]	0.214
English	TIPI (C.)	Trim top decile	4,899	0.429 [0.190, 1.779]	0.134

Note: F.B. = Focused Behaviour; C. = Conscientiousness. ES and percentile confidence intervals are from 1,000 bootstrap replications at noncognitive shift +0.5. The table reports full-sample translog estimates and a support-restricted version that trims the upper tail of the bounded OPS outcome scale. The SDQ result remains stable under trimming; TIPI subgroup uncertainty remains wide.

Table 25: Translog regularity diagnostics: percentage of bootstrap draws in which the estimated production function exhibits economically implausible behaviour (negative marginal products, negative or explosive elasticity of substitution) at sample means. A well-behaved specification should show near-zero violation rates. Shift = +0.5; 500 bootstrap replications.

Sample	Outcome	Noncog	N	% $MP_C < 0$	% $MP_N < 0$	% $ES < 0$	% $ ES > 10$
Full	Maths	SDQ	5,631	0.0	0.0	0.0	0.0
Full	Maths	TIPI	5,631	0.0	0.0	0.2	0.4
Full	English	SDQ	5,631	0.0	0.0	0.0	0.0
Full	English	TIPI	5,631	0.0	0.0	1.9	0.6
Boys	Maths	SDQ	2,667	0.0	0.0	0.0	0.0
Boys	Maths	TIPI	2,667	0.0	0.0	0.2	0.1
Boys	English	SDQ	2,667	0.0	0.0	0.0	0.0
Boys	English	TIPI	2,667	0.0	0.0	4.5	1.0
Girls	Maths	SDQ	2,801	0.0	0.0	0.0	0.0
Girls	Maths	TIPI	2,801	0.0	0.0	11.9	3.9
Girls	English	SDQ	2,801	0.0	0.0	0.0	0.0
Girls	English	TIPI	2,801	0.0	0.1	22.8	3.3

Note: Percentage of bootstrap replications in which the indicated regularity condition is violated at the sample-mean evaluation point. MP_C and MP_N are marginal products of cognitive and noncognitive inputs. ES is the Allen–Uzawa elasticity of substitution. Negative marginal products imply the isoquant slopes upward locally; negative ES implies the curvature condition for substitutability is violated. $|ES| > 10$ flags draws where the ES object is numerically unstable.

6.15 Complementarity across developmental windows

The main translog estimates use Junior Certificate grades (reported at Wave 3) as the achievement outcome, with cognitive and noncognitive inputs measured at Wave 2 (age 13). To test whether the complementarity finding generalises beyond the JC exam and the specific Wave 2 test battery, I re-estimate the translog production function using Leaving Certificate points as the outcome under two alternative specifications.

The first specification uses the same Wave 2 inputs - the PC cognitive composite (Drumcondra Verbal Reasoning, Numerical Ability, and BAS Matrices) and SDQ or TIPI noncognitive measures - but replaces the JC outcome with LC Maths and English points, self-reported at Wave 4 (age 20) for exams sat at ages 17–18. This extends the input–outcome gap from 2–3 years to 4–5 years, spanning the entire secondary-school period.

The second specification uses Wave 3 cognitive tests (semantic fluency, mathematical calculations, and vocabulary - different instruments from those available at Waves 1 and 2) and Wave 3 SDQ or TIPI, with the same LC outcomes. This near-contemporaneous specification tests whether complementarity appears with different cognitive instruments and a different noncognitive measurement occasion.

Table 31 reports the headline results alongside the main JC-based estimates for comparison.

+

Table 26: Sensitivity of ES to evaluation point and noncognitive shift: SDQ (Focused Behaviour) specifications. Unlike the CES, the translog allows the elasticity of substitution to vary across the input space, so the ES estimate can differ depending on where it is evaluated. This table checks whether the $ES < 1$ finding holds at low (25th percentile), mean, and high (75th percentile) input values, and under alternative log-shift values for the noncognitive measure.

Sample / Outcome	Shift	Eval. point	ES (median)	95% CI	% ES < 1
Full Maths	0.5	mean	0.484	[0.374, 0.660]	100.0
Full Maths	0.5	p25	0.486	[0.396, 0.644]	100.0
Full Maths	0.5	p75	0.474	[0.346, 0.667]	100.0
Full Maths	1.0	mean	0.453	[0.350, 0.606]	100.0
Full Maths	1.0	p25	0.456	[0.371, 0.592]	100.0
Full Maths	1.0	p75	0.445	[0.322, 0.610]	100.0
Full English	0.5	mean	0.458	[0.351, 0.660]	100.0
Full English	0.5	p25	0.460	[0.373, 0.638]	100.0
Full English	0.5	p75	0.449	[0.320, 0.672]	100.0
Full English	1.0	mean	0.440	[0.337, 0.603]	100.0
Full English	1.0	p25	0.438	[0.354, 0.581]	100.0
Full English	1.0	p75	0.432	[0.308, 0.615]	100.0
Boys Maths	0.5	mean	0.465	[0.313, 0.822]	99.2
Boys Maths	0.5	p25	0.475	[0.350, 0.814]	99.2
Boys Maths	0.5	p75	0.458	[0.276, 0.836]	99.2
Boys Maths	1.0	mean	0.427	[0.295, 0.746]	99.2
Boys Maths	1.0	p25	0.442	[0.331, 0.737]	99.2
Boys Maths	1.0	p75	0.418	[0.257, 0.757]	99.2
Boys English	0.5	mean	0.402	[0.278, 0.657]	99.8
Boys English	0.5	p25	0.418	[0.317, 0.649]	99.8
Boys English	0.5	p75	0.393	[0.241, 0.674]	99.8
Boys English	1.0	mean	0.386	[0.264, 0.626]	100.0
Boys English	1.0	p25	0.401	[0.301, 0.617]	100.0
Boys English	1.0	p75	0.377	[0.229, 0.640]	100.0
Girls Maths	0.5	mean	0.494	[0.348, 0.817]	99.3
Girls Maths	0.5	p25	0.511	[0.384, 0.808]	99.3
Girls Maths	0.5	p75	0.501	[0.336, 0.836]	99.3
Girls Maths	1.0	mean	0.467	[0.342, 0.764]	99.9
Girls Maths	1.0	p25	0.483	[0.379, 0.759]	99.9
Girls Maths	1.0	p75	0.478	[0.334, 0.779]	99.9
Girls English	0.5	mean	0.374	[0.225, 0.719]	99.7
Girls English	0.5	p25	0.405	[0.288, 0.711]	99.7
Girls English	0.5	p75	0.369	[0.181, 0.746]	99.7
Girls English	1.0	mean	0.363	[0.234, 0.691]	99.4
Girls English	1.0	p25	0.392	[0.289, 0.680]	99.4
Girls English	1.0	p75	0.364	[0.199, 0.728]	99.4

Note: SDQ Focused Behaviour is the primary noncognitive measure. “mean” evaluates ES at sample means of both inputs; “p25” and “p75” evaluate at the 25th and 75th percentiles of both inputs simultaneously. 500 bootstrap replications; percentile CIs.

Table 27: Oster (2019) bounds for omitted variable bias. For each cognitive–noncognitive interaction, δ^* reports the proportional degree of selection on unobservables (relative to observables) required to drive the coefficient to zero. $\delta^* > 1$ indicates that unobservables would need to be more correlated with the interaction than the full set of included controls. $\beta^*(\delta = 1)$ is the bias-adjusted coefficient under equal selection. $R_{\max}^2 = 1.3 \times R_{\text{full}}^2$ (Oster’s recommended heuristic). Full model includes W2 cognitive (PC), noncognitive subscales, SES/school controls, Wave 1 baselines, and all cognition \times noncognition interactions (HC1 SEs).

Outcome	Interaction	$\hat{\beta}$	SE	δ^*	$\beta^*(\delta=1)$	$[\beta, \beta^*]$	Robust
Maths	emot \times Cog	-0.0172	(0.0162)	0.58	0.0127	[-0.0172, 0.0127]	
Maths	cond \times Cog	0.0088	(0.0176)	-0.12	0.0835	[0.0088, 0.0835]	Yes [†]
Maths	hyper \times Cog	-0.0517**	(0.0168)	-6.56	-0.0596	[-0.0596, -0.0517]	Yes [†]
Maths	peer \times Cog	-0.0097	(0.0145)	0.22	0.0340	[-0.0097, 0.0340]	
English	emot \times Cog	-0.0072	(0.0147)	0.37	0.0125	[-0.0072, 0.0125]	
English	cond \times Cog	0.0196	(0.0171)	-0.37	0.0731	[0.0196, 0.0731]	Yes [†]
English	hyper \times Cog	-0.0620***	(0.0161)	4.84	-0.0492	[-0.0620, -0.0492]	Yes
English	peer \times Cog	-0.0198	(0.0141)	0.72	0.0075	[-0.0198, 0.0075]	
Maths	A \times Cog	-0.0119	(0.0162)	0.31	0.0260	[-0.0119, 0.0260]	
Maths	C \times Cog	-0.0144	(0.0154)	0.45	0.0180	[-0.0144, 0.0180]	
Maths	ES \times Cog	-0.0614***	(0.0154)	-1.74	-0.0967	[-0.0967, -0.0614]	Yes [†]
Maths	Ex \times Cog	0.0036	(0.0150)	-0.28	0.0167	[0.0036, 0.0167]	Yes [†]
Maths	O \times Cog	0.0084	(0.0155)	2.89	0.0055	[0.0055, 0.0084]	Yes
English	A \times Cog	-0.0094	(0.0145)	0.84	0.0018	[-0.0094, 0.0018]	
English	C \times Cog	-0.0149	(0.0140)	1.02	-0.0002	[-0.0149, -0.0002]	Yes
English	ES \times Cog	-0.0201	(0.0142)	1.81	-0.0090	[-0.0201, -0.0090]	Yes
English	Ex \times Cog	-0.0013	(0.0131)	0.08	0.0152	[-0.0013, 0.0152]	
English	O \times Cog	-0.0230 ⁺	(0.0135)	4.57	-0.0180	[-0.0230, -0.0180]	Yes

Note: δ^* is the proportional selection on unobservables (relative to observables) needed to drive $\hat{\beta} \rightarrow 0$ (Oster, 2019). Positive $\delta^* > 1$: unobservables would need to exceed observables in explanatory power. Negative δ^* : adding controls moves the coefficient *away* from zero (bias-amplifying), marked [†]. $\beta^*(\delta = 1)$ is the identified-set bound under equal selection. $R_{\max}^2 = 1.3 \times R_{\text{full}}^2$. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 28: Oster (2019) bounds: sensitivity to R_{\max}^2 . Same specifications as Table 27 but with $R_{\max}^2 = \min(2.2 \times R_{\text{full}}^2, 1)$, a more conservative upper bound on attainable R^2 .

Outcome	Interaction	$\hat{\beta}$	R_{\max}^2	δ^*	$\beta^*(\delta=1)$	Robust
Maths	emot \times Cog	-0.0172	1.000	0.19	0.0713	
Maths	cond \times Cog	0.0088	1.000	-0.04	0.2298	Yes [†]
Maths	hyper \times Cog	-0.0517	1.000	-2.22	-0.0750	Yes [†]
Maths	peer \times Cog	-0.0097	1.000	0.07	0.1198	
English	emot \times Cog	-0.0072	0.873	0.09	0.0713	
English	cond \times Cog	0.0196	0.873	-0.09	0.2336	Yes [†]
English	hyper \times Cog	-0.0620	0.873	1.21	-0.0107	Yes
English	peer \times Cog	-0.0198	0.873	0.18	0.0896	
Maths	Agreeable W2 PCG \times Cog	-0.0119	1.000	0.10	0.1026	
Maths	Conscientious W2 PCG \times Cog	-0.0144	1.000	0.15	0.0834	
Maths	Emo Stability W2 PCG \times Cog	-0.0614	1.000	-0.58	-0.1680	Yes [†]
Maths	Extravert W2 PCG \times Cog	0.0036	1.000	-0.09	0.0432	Yes [†]
Maths	Openness W2 PCG \times Cog	0.0084	1.000	0.96	-0.0004	
English	Agreeable W2 PCG \times Cog	-0.0094	0.860	0.21	0.0354	
English	Conscientious W2 PCG \times Cog	-0.0149	0.860	0.25	0.0438	
English	Emo Stability W2 PCG \times Cog	-0.0201	0.860	0.45	0.0242	
English	Extravert W2 PCG \times Cog	-0.0013	0.860	0.02	0.0644	
English	Openness W2 PCG \times Cog	-0.0230	0.860	1.14	-0.0029	Yes

Note: See Table 27 for definitions. [†]Bias-amplifying: controls move coefficient away from zero.

Three patterns emerge from Table 31. First, SDQ-based complementarity ($ES < 1$) is robust to the choice of outcome: the W2→LC English specification produces $ES = 0.46$, nearly identical to the main JC-based estimate, and significantly below one. Second, there is a developmental gradient in Maths: ES rises from 0.48 (W2→JC) to 0.86 (W2→LC) to 1.46 (W3→LC) as inputs are measured closer to the outcome. One interpretation is that early skill interactions are more tightly complementary because they operate through cumulative developmental channels, while contemporaneous measures capture a mix of substitution possibilities that loosen the constraint. Third, TIPI-based estimates are more precise in the LC specifications than in the JC-based main analysis - particularly for English, where both W2→LC and W3→LC produce ES significantly below one - but remain less stable than SDQ across subjects and subgroups. The overall conclusion is that the main complementarity finding is not an artefact of the Junior Certificate outcome, the Wave 2 cognitive test battery, or the specific developmental timing of the main specification.

Table 29: Design-effect adjusted standard errors. Columns report the HC1 coefficient and standard error alongside the cluster-adjusted SE under the conservative scenario ($\bar{m} = 8$, $\rho = 0.15$, $DEFF = 2.05$). This inflates SEs by a factor of $\sqrt{2.05} \approx 1.43$, simulating the effect of school-level clustering without requiring school identifiers. Full model includes W2 cognitive (PC), noncognitive subscales, SES/school controls, Wave 1 baselines, and all cognition \times noncognition interactions.

Outcome	Interaction	$\hat{\beta}$	SE _{HC1}	SE _{adj}	t_{adj}
Maths	emot \times Cog	-0.0172	0.0162	0.0231	-0.74
Maths	cond \times Cog	0.0088	0.0176	0.0251	0.35
Maths	hyper \times Cog	-0.0517	0.0168	0.0240	-2.15
Maths	peer \times Cog	-0.0097	0.0145	0.0208	-0.47
English	emot \times Cog	-0.0072	0.0147	0.0211	-0.34
English	cond \times Cog	0.0196	0.0171	0.0245	0.80
English	hyper \times Cog	-0.0620	0.0161	0.0230	-2.69
English	peer \times Cog	-0.0198	0.0141	0.0203	-0.98
Maths	Agreeable W2 PCG \times Cog	-0.0119	0.0162	0.0231	-0.52
Maths	Conscientious W2 PCG \times Cog	-0.0144	0.0154	0.0221	-0.65
Maths	Emo Stability W2 PCG \times Cog	-0.0614	0.0154	0.0221	-2.78
Maths	Extravert W2 PCG \times Cog	0.0036	0.0150	0.0214	0.17
Maths	Openness W2 PCG \times Cog	0.0084	0.0155	0.0222	0.38
English	Agreeable W2 PCG \times Cog	-0.0094	0.0145	0.0208	-0.45
English	Conscientious W2 PCG \times Cog	-0.0149	0.0140	0.0200	-0.74
English	Emo Stability W2 PCG \times Cog	-0.0201	0.0142	0.0204	-0.98
English	Extravert W2 PCG \times Cog	-0.0013	0.0131	0.0187	-0.07
English	Openness W2 PCG \times Cog	-0.0230	0.0135	0.0193	-1.19

Note: Design effect $DEFF = 1 + (\bar{m} - 1)\rho$ where \bar{m} is the average cluster size and ρ is the intra-class correlation. $SE_{adj} = SE_{HC1} \times \sqrt{DEFF}$. Parameters: $\bar{m} = 8$ students per school (GUI sampling estimate), $\rho = 0.15$ (OECD PISA 2015 between-school variance share for Ireland). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 30: Design-effect sensitivity: SDQ Focused Behaviour \times Cognition (primary noncognitive interaction) across intra-class correlation values. $\bar{m} = 8$. The interaction remains at least marginally significant ($p < 0.10$) for Maths across all ICC scenarios tested.

Outcome	ρ	DEFF	$\hat{\beta}$	SE _{adj}	t_{adj}
Maths	0.05	1.35	-0.0517	0.0195	-2.65
Maths	0.10	1.70	-0.0517	0.0219	-2.36
Maths	0.15	2.05	-0.0517	0.0240	-2.15
Maths	0.20	2.40	-0.0517	0.0260	-1.99
English	0.05	1.35	-0.0620	0.0187	-3.32
English	0.10	1.70	-0.0620	0.0210	-2.96
English	0.15	2.05	-0.0620	0.0230	-2.69
English	0.20	2.40	-0.0620	0.0249	-2.49

Note: Design effect $DEFF = 1 + (\bar{m} - 1)\rho$. $\bar{m} = 8$. $SE_{adj} = SE_{HC1} \times \sqrt{DEFF}$. Full model: Model 4 with Wave 1 baselines. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 31: Translog elasticity of substitution across outcome measures and input timing. The main specification uses Wave 2 inputs (age 13) and Junior Certificate outcomes (age 15). The W2→LC specification uses the same Wave 2 inputs but Leaving Certificate outcomes (age 17/18). The W3→LC specification uses Wave 3 inputs (age 17/18, different cognitive test battery) and LC outcomes.

Specification	Outcome × Noncog	<i>N</i>	ES	95% CI	<i>p</i> (ES = 1)
<i>Panel A: SDQ Focused Behaviour (primary)</i>					
W2→JC (main)	Maths × SDQ	5,631	0.482	[0.378, 0.673]	0.000
W2→LC	Maths × SDQ	3,722	0.864	[0.659, 1.271]	0.420
W3→LC	Maths × SDQ	3,967	1.462	[0.934, 2.618]	0.104
W2→JC (main)	English × SDQ	5,631	0.464	[0.349, 0.690]	0.000
W2→LC	English × SDQ	3,882	0.462	[0.370, 0.595]	0.000
W3→LC	English × SDQ	4,157	0.800	[0.586, 1.168]	0.236
<i>Panel B: TIPI Conscientiousness (confirmatory)</i>					
W2→JC (main)	Maths × TIPI	5,631	0.569	[0.354, 1.701]	0.178
W2→LC	Maths × TIPI	3,722	0.645	[0.486, 1.120]	0.109
W3→LC	Maths × TIPI	3,964	1.176	[0.738, 2.538]	0.557
W2→JC (main)	English × TIPI	5,631	0.514	[0.226, 1.914]	0.214
W2→LC	English × TIPI	3,882	0.420	[0.327, 0.681]	0.006
W3→LC	English × TIPI	4,154	0.550	[0.396, 0.855]	0.014

Note: W2→JC: Wave 2 cognitive composite (PC from Drumcondra + BAS) and Wave 2 noncognitive measures predicting Junior Certificate OPS points. W2→LC: same Wave 2 inputs predicting Leaving Certificate points (self-reported at Wave 4). W3→LC: Wave 3 cognitive composite (standardised average of semantic fluency, mathematical calculations, and vocabulary tests) and Wave 3 noncognitive measures predicting LC points. SDQ is inverted so higher values indicate more positive traits (Focused Behaviour). Translog shift = 0.5. Bootstrap percentile 95% CIs from 1,000 replications. *p*(ES = 1) is a two-sided bootstrap test.