The Origins of the STEM Gender Gap: Evidence from Secondary School Subject Selection

Dáire Crotty[1] Ciarán Murphy[2]

[1]Rutgers University–Camden [2]University College Dublin

Growing Up in Ireland Conference, November 2026

Motivation

1. Secondary choices are understudied

- Secondary subjects are a gateway to STEM study *(Chevalier, 2011; Card & Payne, 2021).*
- Practical subjects predict college STEM enrolment *(Delaney & Devereux, 2019).*
- Little is known about what drives these choices.

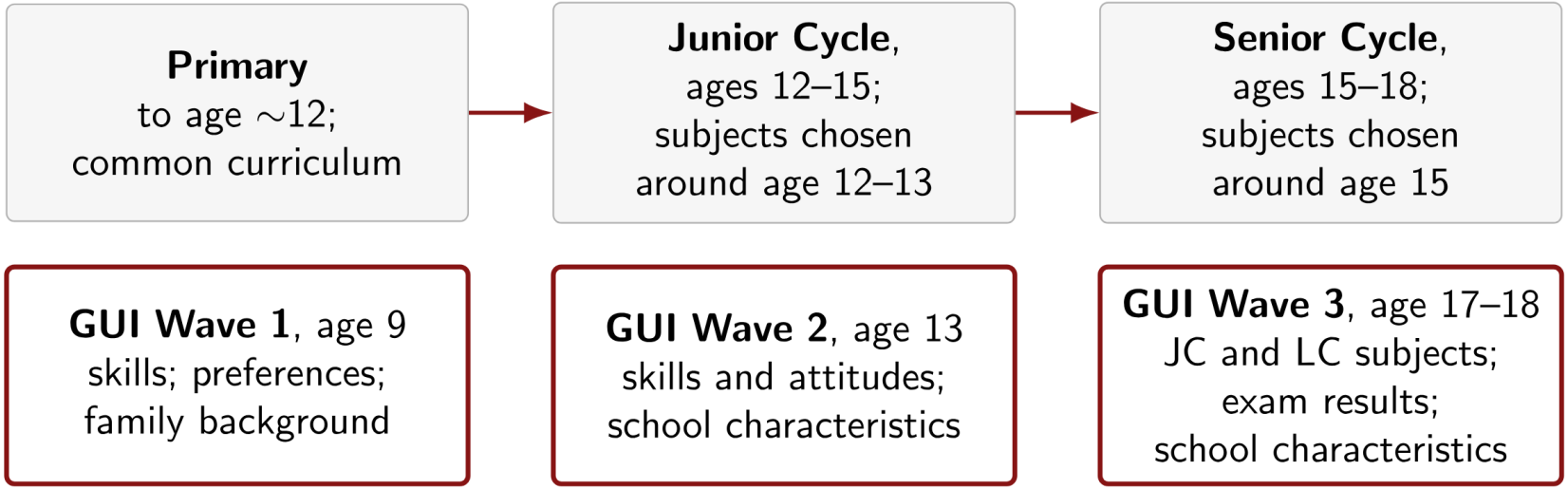
2. STEM is not one thing

- Gender gaps in STEM are large in higher education and the labour market.
- But "STEM" bundles very different subjects.
- Which subjects drive the gap?

Which STEM subjects diverge by gender, and when does the divergence begin?

Institutional Setting and Data

- Growing Up in Ireland, Cohort '98. Analytical sample: 3,888 students



- Leaving Cert subjects, chosen around age 15, shape access to STEM in college: what explains the gender gaps in these choices?

Methodology: Regression and Decomposition

1. Raw gender gap

$$Y_i = \alpha + \beta^{\text{raw}} \, \text{Male}_i + \varepsilon_i$$

2. Adding mechanism groups

$$Y_i = \alpha + \beta^{\text{full}} \, \text{Male}_i + \mathbf{X}_i' \boldsymbol{\gamma} + \varepsilon_i$$

3. Gelbach (2016) decomposition

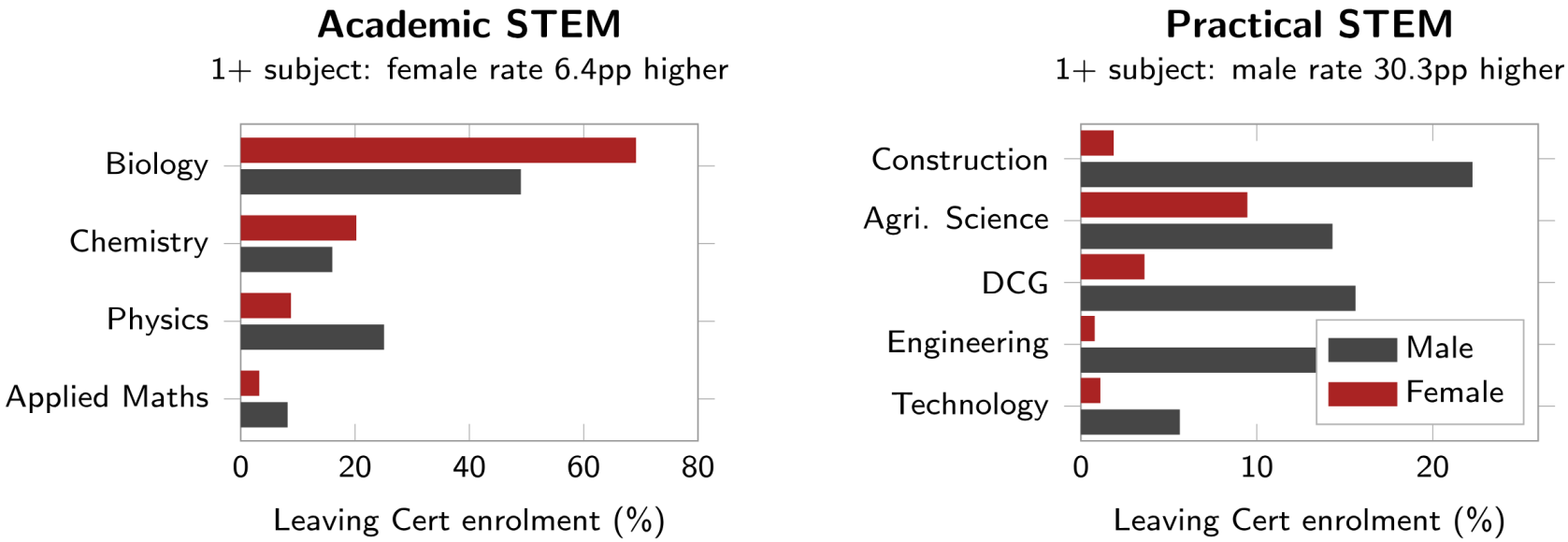
$$\beta^{\text{raw}} - \beta^{\text{full}} = \sum_g \Delta^g$$

$Y_i = 1$ if student i takes at least one practical STEM subject at Leaving Cert. Linear probability model; survey weights; standard errors clustered by school.

Mechanism Groups and Variables

Mechanism group	Explanation in the literature	Variables
STEM-readiness	Earlier exposure shapes later specialisation *Fricke et al. (2018); Gottfried & Bozick (2016)*	Each optional Junior Cycle subject taken; Transition Year
Skills and comparative advantage	Ability, and relative strengths across different subject areas *Speer (2017); Goulas et al. (2024); Buser et al. (2017)*	Drumcondra reading and numerical scores (age 9); JC maths and English grades; number of higher-level JC subjects; behaviour (SDQ); conscientiousness; academic self-concept
School factors and peers	Subject offerings, peers, single-sex schooling *Smith & Hannan (2006); Mouganie & Wang (2020); Schneeweis & Zweimüller (2012); Calkins et al. (2023)*	Whether the school offers each subject; DEIS; fee-paying; single-sex or mixed
Preferences and tastes	Interests, self-confidence, aspirations *Stokking (2000); Kuhn & Wolter (2022); Mackenzie et al. (2021)*	Liking maths, reading and school; favourite subject; perceived difficulty of maths and English; occupational aspirations; educational expectations
Background	Family and individual controls	Caregiver education and country of birth; books in the home; disability

Gender Gaps in STEM Enrolment, by Subject Type



Practical subjects strongly predict college STEM enrolment (Delaney & Devereux, 2019), yet are rarely studied.

Results: Leaving Certificate Practical STEM

	Percentage points	(SE)	% of raw gap
Raw male–female gap	30.3***	(2.3)	100
Explained by:			
STEM-readiness (earlier JC choices)	17.4***	(2.3)	57
Preferences / tastes	5.8***	(1.4)	19
Skills / comparative advantage	0.04	(0.5)	<1
School factors / peers	0.01	(1.8)	<1
Background	−0.05	(0.3)	<1
Remaining gap (full-model male coefficient)	7.1**	(3.4)	23

Outcome: any practical STEM subject at LC. N = 3,888. Survey weights; SEs clustered by school. *** $p < 0.01$, ** $p < 0.05$.

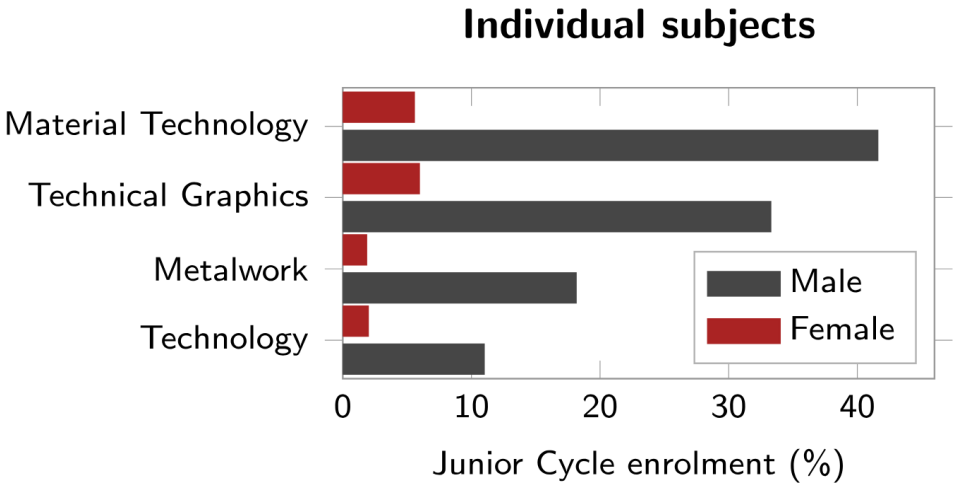
Results: Leaving Certificate Practical STEM

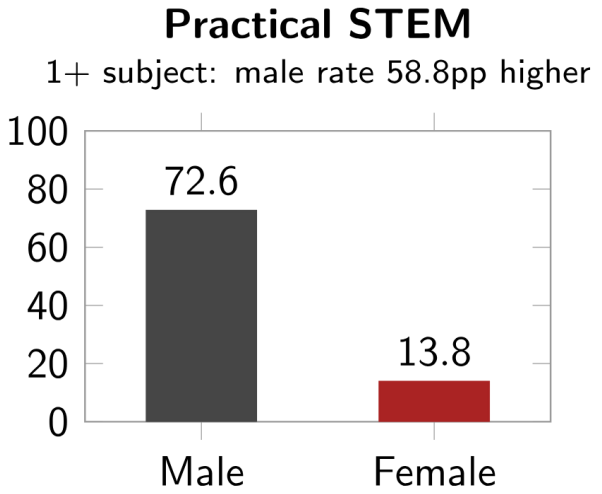
	Percentage points	(SE)	% of raw gap
Raw male–female gap	30.3***	(2.3)	100
Explained by:			
STEM-readiness (earlier JC choices)	17.4***	(2.3)	57
Preferences / tastes	5.8***	(1.4)	19
Skills / comparative advantage	0.04	(0.5)	<1
School factors / peers	0.01	(1.8)	<1
Background	−0.05	(0.3)	<1
Remaining gap (full-model male coefficient)	7.1**	(3.4)	23

Outcome: any practical STEM subject at LC. N = 3,888. Survey weights; SEs clustered by school. *** $p < 0.01$, ** $p < 0.05$.

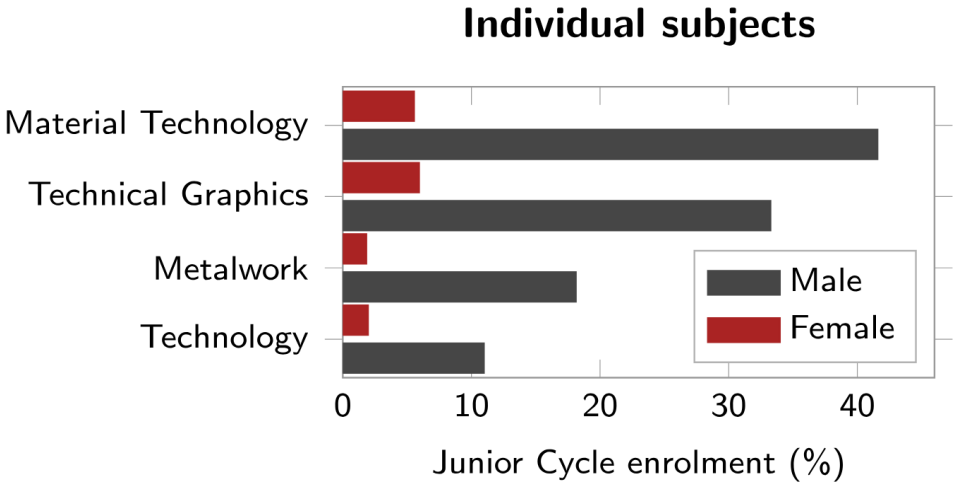
So where does the Junior Cycle gap come from?

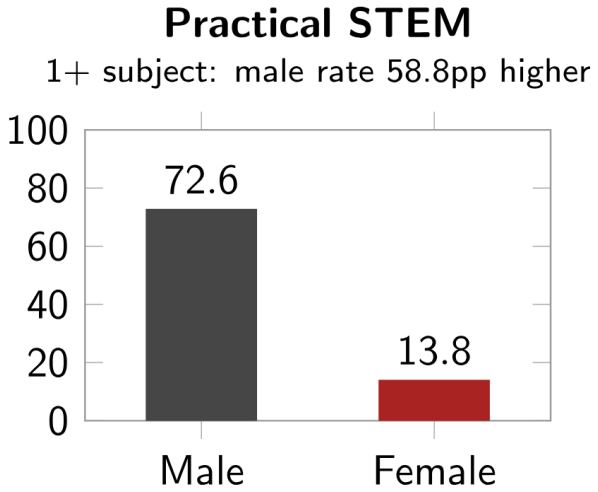
Gender Gaps in Junior Cycle Practical STEM





Gender Gaps in Junior Cycle Practical STEM





- Decompose 58.8pp gap, using only age-9 information (before subject and school choice)
- STEM-readiness and school factors drop out; skills, preferences, background remain

Results: Junior Cycle Practical STEM

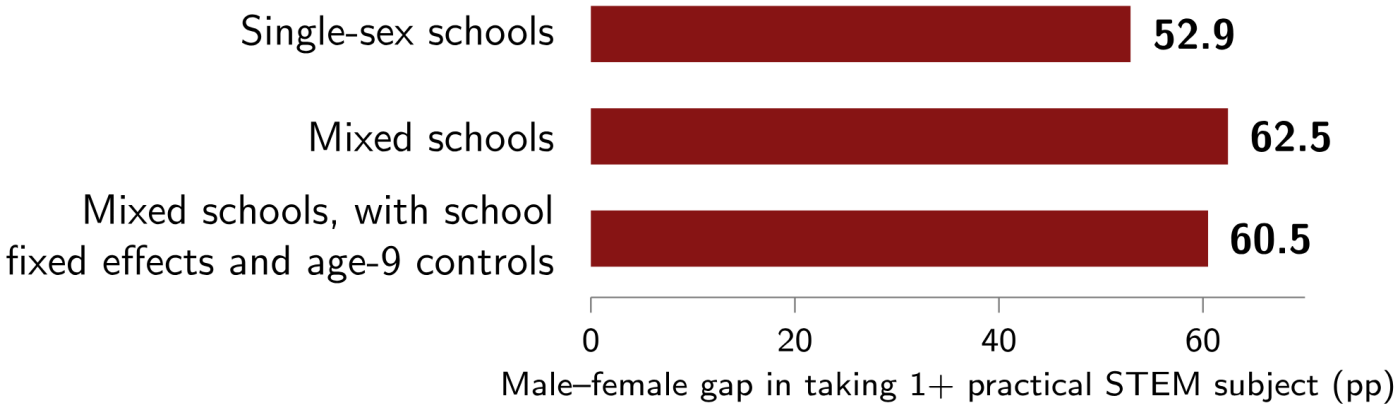
	Percentage points	(SE)	% of raw gap
Raw male–female gap	58.8***	(3.4)	100
Explained by:			
Preferences / tastes	2.2**	(1.0)	4
Skills / comparative advantage	−0.2	(0.4)	≈0
Background	−0.2	(0.4)	≈0
Total explained	1.8*	(1.1)	3
Remaining gap (full-model male coefficient)	57.0***	(3.2)	97

Outcome: any practical STEM subject at JC. N = 3,888. Survey weights; SEs clustered by school. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Age-9 characteristics explain 1.8pp of the 58.8pp gender gap.

Robustness: School Sorting and Subject Availability

- Around half of Irish students attend single-sex schools, which may offer different subjects
- If availability drove the gap, it should be smaller in mixed schools (same subject menu)



- The gap is larger in mixed schools, and persists within schools

Robustness: Machine-Learning Decomposition

Not explained by schools or variables, but could it reflect the linear model or variable choice?

Fully flexible model

gradient-boosted trees (XGBoost) nonlinearities and interactions

66 extra variables

→

SHAP values

split the predicted gap additively across our mechanism groups, even for a highly nonlinear model

→

Same conclusions

earlier choices matter at Leaving Cert; age-9 observables explain little at Junior Cycle

The results are not driven by the model's parametric assumptions or by our choice of variables.

Conclusion and Policy Implications

1. **The gap is in practical subjects, and opens early**
   - Gaps are concentrated in practical subjects
   - Gaps are largely predictable from choices made at age 12–13
2. **Early gap not explained by subject availability, skills or observed preferences**
   - Points to factors we don't observe:
     - gender norms, stereotypes and attitudes
     - differential encouragement from parents, teachers and peers
3. **Future research**
   - Attitudes about gender and specific subjects, measured before subject choice
   - Interest in practical subjects, tracked between ages 9 and 13

Thank You

Questions and feedback welcome!

Ciarán Murphy
School of Economics, University College Dublin

ciaran.murphy10@ucdconnect.ie

LC STEM Enrolment by Gender

Subject	Male (%)	Female (%)	Gap (pp)
Academic STEM			
Biology	48.88	69.03	-20.15
Chemistry	15.89	20.09	-4.20
Physics	24.94	8.66	16.28
Applied Maths	8.07	3.11	4.96
Any academic STEM	**70.46**	**76.86**	**-6.40**
Practical STEM			
Construction Studies	22.22	1.82	20.40
Agricultural Science	14.26	9.42	4.84
Design & Communication Graphics	15.57	3.57	12.00
Engineering	14.21	0.74	13.47
Technology	5.58	1.06	4.52
Any practical STEM	**45.96**	**15.63**	**30.33**

JC Practical STEM Enrolment by Gender

Subject	Male (%)	Female (%)	Gap (pp)
Material Technology	41.57	5.54	36.03
Technical Graphics	33.25	5.93	27.32
Metalwork	18.12	1.83	16.29
Technology	10.97	1.96	9.01
Any JC practical STEM	**72.55**	**13.78**	**58.77**

Mixed Schools with School Fixed Effects

	LC Practical STEM	JC Practical STEM
Baseline male coefficient + school FE	0.336*** (0.027)	0.625*** (0.026)
Full model + school FE	0.086** (0.037)	0.605*** (0.031)
Total change	0.250***	0.020*
STEM-readiness	0.218***	–
Skills / comparative advantage	-0.004	0.001
Preferences / tastes	0.044**	0.018
Background characteristics	-0.008	-0.0002
Observations	1,988	1,988

School fixed effects absorb subject availability and other school-level characteristics.

Machine Learning: SHAP Decomposition

	LC Practical STEM	JC Practical STEM
Predicted gap (test data)	31.15	50.72
Contribution to predicted gap (pp):		
STEM-readiness	14.33	–
School factors / peer effects	1.94	0.61
Preferences / tastes	5.45	2.82
Skills / comparative advantage	1.00	0.63
Background characteristics	0.17	0.20
Gender (residual)	8.26	46.46

70/30 train–test split; expanded models add 66 variables to the baseline set.
Test AUC: 0.80 (LC), 0.86 (JC).