What Determines Post-Compulsory Academic Studies? Evidence from the Longitudinal Survey of Young People in England

Collier, William; Valbuena, Javier; Zhu, Yu

Published in:
Applied Economics Letters

DOI:
10.1080/13504851.2017.1352069

Publication date:
2018

Document Version
Peer reviewed version

Word count: 1,986 (2,212 including cover page)

What Determines Post-Compulsory Academic Studies? Evidence from the Longitudinal Survey of Young People in England

William Collier[a], Javier Valbuena[b] and Yu Zhu[c]

[a] School of Economics, University of Kent, Canterbury, CT2 7NP. E-mail: W.J.Collier@kent.ac.uk
[b] Corresponding author: Department of Economics, University of Girona, Campus Montilivi, 17003, Girona. E-mail: valbuena.javier@gmail.com
[c] Economic Studies, University of Dundee, 3 Perth Road, Dundee, DD1 4HN, United Kingdom. E-mail: yuzhu@dundee.ac.uk

ABSTRACT

We show that educational attainments at the end of the compulsory schooling stage are powerful predictors for post-compulsory educational choices in England. In particular, the single academic success indicator of achieving the Government's *gold standard in GCSE*, is able to predict virtually all the observed incidences of post-compulsory studies for academic qualifications. Notwithstanding, Two Stage Least Squares estimation which exploits variations in school starting age induced by school entry rules suggests that the least-squares effect of achieving the *gold standard in GCSEs* on studying for academic qualifications is due to ability bias or reverse causation.

Keywords: Educational choice, Two Stage Least Squares, Gold standard in GCSEs, Relative school starting age

JEL Classification: I21, J24, P36

Acknowledgements: We wish to thank Antonio Di Paolo, Colm Harmon, Jagjit Chadha and participants at the WPEG conference, the First Lisbon Research Workshop on Economics and Econometrics of Education and the XIX Meeting of the Economics of Education Association for their helpful comments. All errors and omissions remain the responsibility of the authors.

This is an Accepted Manuscript of an article published by Taylor & Francis in *Applied Economics Letters* on 13 July 2017, available online: http://www.tandfonline.com/10.1080/13504851.2017.1352069.

1. Introduction

It is well established that the returns to academic qualifications are higher than those to vocational qualifications at the same nominal level in the UK (see e.g. Dearden *et al*. 2002, McIntosh 2006 and Dickerson 2008). We focus on the determinants of studying for any academic qualifications beyond compulsory education, using a unique dataset which is rich in both family background and attainment in education. We show that the single indicator of achieving the Government's *gold standard in the General Certificate of Secondary Education (GCSE)* is able to predict virtually all the observed aggregate proportion of young people studying for academic qualifications.[1]

We also investigate the extent to which the impact of initial academic success on post-compulsory educational choices reflects a causal relationship. Two Stage Least Squares (2SLS) estimates which exploit variations in school starting age by month of birth induced by school entry rules suggest that the effect of achieving the *gold standard in GCSEs* on studying for academic qualifications is due to individual heterogeneity (ability bias) or simultaneity bias (reverse causation).

2. Data and set-up of the analysis

We use the Longitudinal Study of Young People in England (LSYPE), a panel study of young people who were studying in Year 9 in England in the school year 2003-2004.

We select UK-born young people who provide full interviews, together with their mothers who are aged 60 or below, across the first 4 waves of the LYSPE data.[2] This results in a final sample of 9,190 young people, of which 4,570 (49.7%) are boys for our analysis.

Table 1 presents summary statistics by gender for all variables used in our empirical analyses. The first thing to note is a significant gender gap in favour of girls in our outcome variable, with 61% of girls and 52% of boys studying for academic qualifications beyond compulsory education.[3] Similarly, there is a significant gender gap in educational attainment at around age 16 in favour of girls: 57% of girls achieve the critical benchmark of 5 or more

[1] GCSE stands for General Certificate of Secondary Education. Students achieve the "gold standard" in GCSE if they achieve five or more A* to C passes, including English and Mathematics.

[2] Young persons born before the 1st September 1989 or after the 31st August 1990 are excluded from the sample on the basis that they violate the school entry rule for the Wave 1 cohort.

[3] Academic qualifications include A Levels and its component units, AVCEs (Advanced Vocational Certificate in Education), or GCSEs. There are only 51 young people taking AVCEs, which are full-time education based at schools or colleges, unlike traditional vocational routes such as apprenticeships. Of the 351 young people studying for GCSEs post-compulsory schooling, only 8% have achieved the *gold standard*, suggesting many of them are retaking subject (or retaking examinations) to improve their grades.

GCSEs at Grades A*-C including Maths and English, whilst only 49% of boys manage to reach the same standard. Given the significant gender gap in the key variables, we estimate gender specific regressions.

3. Empirical results

3.1 Linear Probability Model (LPM) of the effect of achieving the *gold standard*

It is conventional wisdom that the single most important predictor for studying for academic qualifications beyond age 16 in England is represented by achievement of the *gold standard in GCSEs* (see e.g. Nuffield Foundation 2009).

The first two columns of Table 2 assess the importance of achieving the *gold standard* on the probability of studying for academic qualifications for boys and girls separately, controlling for family characteristics.[4] It is striking that the partial effect of achieving the *gold standard* is able to predict virtually all of the observed mean probability of studying for post-compulsory academic qualifications for boys and almost 80% of the probability for girls. In terms of the total variation in the dependent variable, our parsimonious specification can explain 36% for boys and 32% for girls.

3.2 Causal effect

In columns 3 and 4, we present the corresponding Two Stage Least Squares (2SLS) estimates. The causal effect of achieving the *gold standard* is identified through exogenous variation in the relative school starting age (SSA) by month of birth within the same school cohort group, induced by the school entry policy in England. Under a single-entry-point system, a child born on the 1st September 1989 will be the oldest in this school cohort while another child born on the 31st August 1990 will be the youngest.[5]

Drawing on 18 research studies for a range of countries, Sharp *et al*. (2009) conclude that there is overwhelming evidence of statistically significant effects for relative age. Pupils who are younger in the year group fare less well in attainment tests, commonly measured by test scores in maths, reading and writing. For recent UK evidence, see e.g. Crawford *et al*. (2007).

[4] We omit the type of school indicator and parental aspirations from Wave 1, for fear of endogeneity problems. The mother's employment status and family income variables are excluded for similar reasons.

[5] Crawford *et al*. (2007) estimate that around one half of all children born between 1997 and 1999 started school in a Local Education Authority where a single-entry-point system was in operation.

We expect some noise in the actual SSA (which we do not observe in our data), due to the fact that different school entry rules are in operation in different LEAs.[6] However, what matters for identification purposes is whether month of birth is *statistically* correlated with the probability of achieving *gold standard* while having no direct impact on the post-compulsory educational choice. Our data clearly shows that a September-born is more likely than her August-born counterpart to pass the threshold. This implies that on average, the chance of reaching the *gold standard* for both boys and girls is increasing in the predicted SSA (using the single-entry-point rule).

By and large, the family background variables maintain their signs and statistical significances when we endogenise the *gold standard* indicator, although the sizes of the effects appear to be larger under the 2SLS specification.

In contrast, the 2SLS estimates for achieving the *gold standard* are 54% lower than their LPM counterparts for boys and 67% lower for girls respectively. Moreover, none of the 2SLS estimates are close to being statistically significant ($p>0.35$). This implies that the effect of achieving the *gold standard* on choosing the academic pathway is driven by individual heterogeneity (ability) or reverse causation.

Controlling for family background, a 10-month difference in SSA results in 7 and 11 percentage points difference in the probability of achieving the *gold standard* for boys and girls, respectively. Furthermore, the Cragg-Donald Wald F-statistics for the excluded restrictions are well above the recommended threshold of 10 in both models, meaning we do not have a weak-instrument problem. Indeed, the F-statistics are above the critical value for 15% and 10% relative bias for boys and for girls respectively, implying that our 2SLS estimates have been successful in removing most of the bias in the LPM estimates.[7]

4. Conclusions

Using a unique dataset which is rich in both family background and attainment in education, we find the single academic success indicator of achieving the Government's *gold standard in GCSE* is able to predict virtually all the observed incidences of post-compulsory studies for academic qualifications.

[6] Admittedly, children exposed to multi-entry-points systems will receive different length of education (up to 2 terms) at the end of the compulsory education stage. This idea has been exploited by Del Bono and Galindo-Rueda (2004) for the UK and Black *et al*. (2011) for Norway.

[7] To check the validity of our findings under the presence of heteroskedasticity, we re-estimated the results in Table 2 with a robust VCE. The robust standard errors and associated diagnostic tests are qualitatively identical.

However, the corresponding 2SLS estimates which exploit variations in school starting age by months of birth induced by school entry rules are statistically insignificant, with $p>0.35$ for both gender. This implies that the effect of achieving the *gold standard* on pursuing the academic pathway is non-causal.

To the extent that post-compulsory educational choices reflect differences in pupils' ability, a substantial and persistent earnings gap between the academic and vocational qualifications at the same notional level would seem unavoidable, despite the government's effort to promote the parity of esteem of the two tracks.

References

Black, S., P. Devereux, and K.G. Salvanes (2011), 'Too Young to Leave the Nest? The Effects of School Starting Age', *Review of Economics and Statistics*, 93(2), 455-467.

Crawford, C., Dearden, L. and C. Meghir (2007), 'When You Are Born Matters: the Impact of Date of Birth on Child Cognitive Outcomes in England', *IFS Commentary*, www.ifs.org.uk/docs/born_matters_report.pdf.

Dearden, L., McIntosh, S., Myck, M. and A. Vignoles (2002), 'The Returns to Academic and Vocational Qualifications in Britain', *Bulletin of Economic Research*, 54(3), 249-274.

Del Bono, D. and F. Galindo-Rueda (2004), 'Do a Few Months of Compulsory Schooling Matter? The Education and Labour Market Impact of School Leaving Rules', *IZA DP* 1233.

Dickerson, A. (2008), 'The Distribution and Returns to Qualifications in the Four Countries of the UK', *Sector Skills Development Agency Research Report*, 21A.

McIntosh, S. (2006), 'Further Analysis of the Returns to Academic and Vocational Qualifications', *Oxford Bulletin of Economics and Statistics, 68*, 225–251.

Nuffield Foundation, (2009), *The Nuffield Review of 14-19 Education*. http://www.nuffield14-19review.org.uk/.

Sharp, C., George, N., Sargent, C., O'Donnell, S. and M. Heron (2009), *International Thematic Probe: The Influence of Relative Age on Learner Attainment and Development*, National Foundation for Educational Research.

Table 1: Summary Statistics for Family Characteristics

Variable Name	Boys	Girls
Studying for any academic qualifications (dep. var.)	0.518 (0.500)	0.611 (0.487)
Gold standard in GCSEs	0.487 (0.500)	0.566 (0.496)
Family Characteristics (measured at age 16, or Wave 4):		
Mother non-white	0.223 (0.416)	0.251 (0.434)
Mother's highest qualification is degree or above	0.122 (0.327)	0.118 (0.323)
Mother's highest qualification is NQF4 but below degree	0.138 (0.345)	0.131 (0.338)
Mother's highest qualification is NQF3	0.135 (0.342)	0.135 (0.342)
Mother's highest qualification is NQF2	0.285 (0.452)	0.294 (0.456)
Mother's highest qualification is NQF1	0.115 (0.319)	0.113 (0.317)
Mother has no qualification (reference category)	0.205 (0.404)	0.211 (0.408)
Mother's highest qualification is vocational	0.309 (0.462)	0.305 (0.460)
Mother Married (reference category)	0.704 (0.457)	0.687 (0.464)
Mother cohabiting	0.059 (0.235)	0.054 (0.227)
Mother is lone parent	0.238 (0.426)	0.259 (0.438)
Indicator for step-family	0.102 (0.302)	0.097 (0.296)
Number of siblings to the Young Person	1.503 (1.167)	1.530 (1.181)
Number of Observations	4570	4620

Notes: Standard errors in parentheses.

Table 2: Comparing LPM with 2SLS, All young people aged 16/17

Studying for any academic qualifications	**LPM**		**2SLS**	
	Boys	**Girls**	**Boys**	**Girls**
Second Stage Results:				
Gold standard in GCSEs	**0.517**	**0.475**	0.240	0.156
	(0.013)	(0.013)	(0.271)	(0.171)
Family Characteristics (measured at age 16 or Wave 4):				
Mother non-white	**0.170**	**0.170**	**0.190**	**0.204**
	(0.016)	(0.015)	(0.025)	(0.025)
Mother's highest qualification is degree+	**0.146**	**0.146**	**0.277**	**0.305**
	(0.023)	(0.023)	(0.131)	(0.089)
Mother's highest qualification is NQF4	**0.137**	**0.129**	**0.261**	**0.269**
	(0.029)	(0.029)	(0.125)	(0.081)
Mother's highest qualification is NQF3	**0.078**	**0.064**	**0.170**	**0.175**
	(0.024)	(0.025)	(0.093)	(0.065)
Mother's highest qualification is NQF2	0.029	**0.051**	**0.082**	**0.124**
	(0.019)	(0.019)	(0.056)	(0.043)
Mother's highest qualification is NQF1	-0.030	-0.002	-0.005	0.026
	(0.023)	(0.023)	(0.034)	(0.029)
Mother's highest qualification is vocational	-0.029	*-0.032*	**-0.061**	**-0.066**
	(0.018)	(0.019)	(0.036)	(0.027)
Mother cohabiting	*-0.046*	**-0.090**	**-0.088**	**-0.112**
	(0.028)	(0.028)	(0.050)	(0.032)
Mother is lone parent	**-0.077**	**-0.087**	**-0.122**	**-0.143**
	(0.014)	(0.014)	(0.046)	(0.034)
Indicator for step-family	**-0.048**	**-0.066**	**-0.070**	**-0.101**
	(0.021)	(0.022)	(0.031)	(0.030)
Number of siblings to the Young Person	**-0.014**	-0.007	**-0.025**	*-0.017*
	(0.005)	(0.005)	(0.011)	(0.008)
R^2	0.356	0.319		
Root MSE	0.403	0.402	0.421	0.428
First Stage Results:				
Relative (school starting) age in months			**-0.007**	**-0.011**
			(0.002)	(0.002)
Diagnostic Tests:				
Cragg-Donald Wald F-stat for excluded Restrictions			**11.17**	**29.76**
(p-value)			(0.001)	(0.000)
Critical Value for 10% relative bias			16.38	
Critical Value for 15% relative bias			8.96	

Notes: Standard errors in parentheses. **Bold** and *italic* cases indicate statistical significance at the 5% and the 10% levels respectively.