The Long-Run Effects of Attending an Elite School: Evidence from the UK (Damon Clark and Emilia Del Bono): Online Appendix

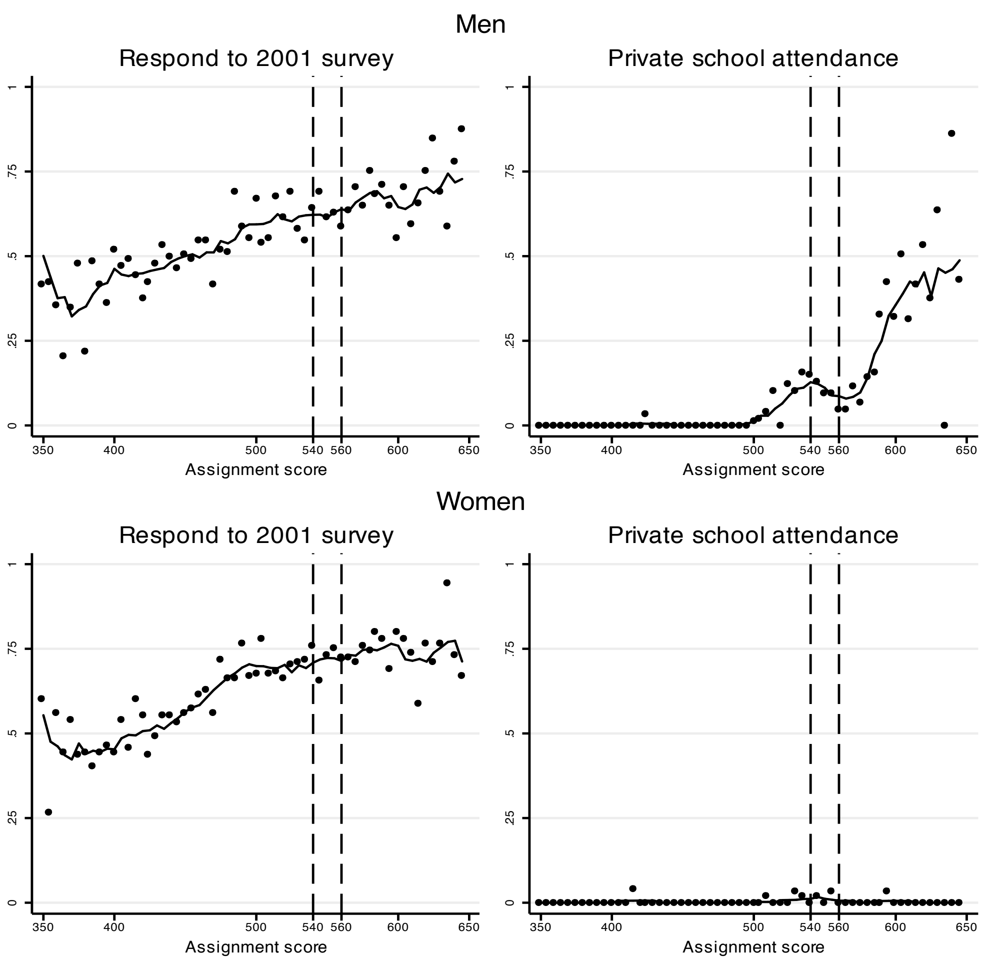
Appendix A: Additional Figures and Tables

Appendix B: Aberdeen Cohort and Labour Force Survey Comparison

Appendix C: A Model of School Quality with Vocational Training

Appendix A: Additional Figures and Tables

Appendix A Figure 1: Assignment scores, survey response and private school attendance



Notes: figures based on "base sample" (see text and Appendix Table 1). Circles represent cell-specifc outcome means. Cells defined over 5-score intervals (350-354, etc.). Vertical lines plotted for cells 540-544 and 560-564.

Appendix A Table 1: sample selection and subsample definitions

	N	
Original sample	12150	
Did not move during 1962-64 (358)		May have left Aberdeen
Not private primary school (255)		Some of these didn't take test
Not RC primary school (299)		Most took test, but stayed in RC school
Not RC private primary school (81)		Same arguments as previous two
Not primary elite school already (504)		Always stay in elite school
Not primary special school (167)		Some of these didn't take test
Not primary school outside Aberdeen (12)		Attend secondary outside Aberdeen
Not missing grade info (47)		Because we want to split by grade
Not missing assignment info (172)		Need to construct IV
Not missing test scores at ages 7, 9 (309)		Need for covariate controls
Not secondary outside Aberdeen (123)		No secondary info
Base sample	9893	
Not in grade 3 in 1962 (2033)		Different assignment procedure
Not missing 2001 survey responses (3151)		Needed for outcomes
Final sample	4709	

Notes: Description of sample selection procedure applied to original data.

Appendix A Table 2: Falisification tests of elite school impacts

	Men			Women		
	OLS	OLS+X	2SLS	OLS	OLS+X	2SLS
Years of education	2.228	0.486	-0.131	2.184	0.230	0.033
	(0.092)	(0.098)	(0.148)	(0.097)	(0.091)	(0.122)
A levels	0.399	0.075	-0.012	0.395	0.034	0.002
	(0.016)	(0.016)	(0.024)	(0.017)	(0.015)	(0.020)
Degree	0.270	0.066	-0.004	0.263	0.027	0.013
	(0.012)	(0.013)	(0.020)	(0.013)	(0.011)	(0.015)
Log Annual Income	0.362	0.037	-0.002	0.345	0.001	-0.028
	(0.016)	(0.015)	(0.024)	(0.017)	(0.014)	(0.021)
Employed	0.053	-0.007	-0.008	0.047	-0.008	-0.016
	(0.004)	(0.004)	(0.006)	(0.003)	(0.004)	(0.005)
Imputed hourly wage	0.288	0.042	-0.031	0.281	0.017	-0.005
	(0.012)	(0.011)	(0.017)	(0.012)	(0.011)	(0.015)
Number of children	-0.095	-0.063	-0.046	-0.116	-0.054	-0.041
	(0.009)	(0.013)	(0.022)	(0.009)	(0.013)	(0.020)
Any children	-0.040	-0.015	-0.001	-0.042	-0.013	-0.006
	(0.003)	(0.004)	(0.007)	(0.003)	(0.004)	(0.006)
Currently married	0.018	-0.005	0.001	0.018	0.000	-0.008
	(0.004)	(0.005)	(0.009)	(0.003)	(0.005)	(0.007)

Notes: see notes to Table 2. The dependent variables in these models are those predicted by a regression of the relevant outcomes on dummies for father's occupation, mother's socio-economic status, school and grade attended in 1962, relative age within the school-grade and third-order polynomials in the age-7 and age-9 test scores.

Appendix A Table 3: Impact of elite school attendance: robustness check

	Men			Women		
	2SLS			2SLS		
	Order 2	Order 3	Order 4	Order 2	Order 3	Order 4
Years of education	1.029	0.925	0.667	0.813	0.793	0.604
	(0.309)	(0.310)	(0.430)	(0.262)	(0.254)	(0.311)
A levels	0.162	0.105	0.051	0.245	0.229	0.122
	(0.061)	(0.064)	(0.090)	(0.054)	(0.052)	(0.067)
Degree	0.153	0.170	0.184	0.024	0.036	0.023
	(0.056)	(0.056)	(0.078)	(0.044)	(0.042)	(0.052)
Log Annual Income	-0.081	-0.115	-0.251	0.138	0.154	0.255
	(0.075)	(0.081)	(0.110)	(0.100)	(0.098)	(0.129)
Employed	-0.042	-0.041	0.039	-0.018	-0.010	0.031
	(0.034)	(0.040)	(0.056)	(0.036)	(0.037)	(0.053)
Imputed hourly wage	0.003	-0.035	-0.136	0.072	0.073	0.083
	(0.054)	(0.056)	(0.080)	(0.051)	(0.050)	(0.061)
Number of children	0.259	0.278	0.259	-0.372	-0.388	-0.585
	(0.171)	(0.176)	(0.256)	(0.124)	(0.128)	(0.169)
Any children	0.020	0.017	-0.027	-0.087	-0.081	-0.140
	(0.060)	(0.060)	(0.087)	(0.043)	(0.042)	(0.054)
Currently married	-0.013	-0.004	-0.073	0.056	0.051	0.030
	(0.062)	(0.063)	(0.089)	(0.047)	(0.048)	(0.065)

Notes: see notes to Table 2. The 2SLS estimates correspond to the "2SLS" estimates in Tables 1-3. Those estimates used a third-order polynomial hence are reproduced in the "Order 3" column. The "Order 2" and "Order 4" columns report estimates obtained with second- and fourth-order polynomials respectively.

Appendix A Table 4: Education impacts by SES

	Low-SES				High-SES			
	OLS	OLS+X	2SLS	2SLS+X	OLS	OLS+X	2SLS	2SLS+X
Years of education	2.951	0.961	0.738	0.732	3.112	1.294	0.653	0.650
	(0.146)	(0.209)	(0.253)	(0.256)	(0.119)	(0.180)	(0.297)	(0.285)
	Control mean=2.05, N=2576				Control mean=2.36, N=2068			
A levels	0.565	0.205	0.164	0.151	0.539	0.251	0.163	0.157
	(0.028)	(0.043)	(0.053)	(0.054)	(0.022)	(0.036)	(0.057)	(0.058)
	Control mean=0.41, N=2576				Control mean=0.46, N=2068			
Degree	0.341	0.098	0.067	0.063	0.398	0.135	0.091	0.075
	(0.025)	(0.036)	(0.044)	(0.043)	(0.022)	(0.029)	(0.053)	(0.050)
	Control mean=0.16, N=2576				Control mean=0.16, N=2068			
Log Annual Income	0.494	0.122	0.149	0.131	0.394	0.143	-0.111	-0.056
	(0.044)	(0.069)	(0.101)	(0.094)	(0.042)	(0.066)	(0.116)	(0.108)
	Control mean=9.67, N=2521				Control mean=9.71, N=2032			
Employed	0.075	0.002	0.026	0.047	0.033	-0.023	-0.058	-0.070
	(0.016)	(0.029)	(0.038)	(0.042)	(0.015)	(0.025)	(0.042)	(0.041)
	Control mean=0.91, N=2576				Control mean=0.89, N=2068			
Imputed hourly wage	0.347	0.016	-0.017	-0.042	0.352	0.106	0.005	0.039
	(0.029)	(0.041)	(0.057)	(0.051)	(0.021)	(0.032)	(0.061)	(0.057)
	Control mean=2.07, N=2284				Control mean=2.13, N=1838			
Number of children	-0.111	-0.103	-0.230	-0.204	-0.116	-0.053	-0.019	0.015
	(0.057)	(0.102)	(0.146)	(0.151)	(0.047)	(0.092)	(0.140)	(0.139)
	Control mean=2.05, N=2575				Control mean=1.89, N=2064			
Any children	-0.050	-0.007	-0.042	-0.028	-0.047	-0.041	-0.055	-0.063
	(0.020)	(0.034)	(0.050)	(0.052)	(0.017)	(0.034)	(0.050)	(0.050)
	Control mean=0.89, N=2574				Control mean=0.88, N=2062			
Currently married	0.026	0.005	0.017	0.027	0.003	0.002	0.016	0.016
	(0.025)	(0.041)	(0.065)	(0.067)	(0.018)	(0.033)	(0.051)	(0.053)
	Control mean=0.79, N=2570				Control mean=0.84, N=2064			

Notes: see notes to Table 2. Low-SES is defined as father's social class 1 through 4 (lower-skilled, semi-skilled, unskilled manual or not working); high-SES is defined as father's social class 5 through 8 (skilled manual occupation that required an apprenticeship, non-manual, intermediate or professional occupation).

Appendix B: Aberdeen Cohort and Labour Force Survey Comparison

Appendix B Table 1: Aberdeen Study versus Labour Force Survey: age left full-time education

	MEN			WOMEN		
	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland
<=15	42.98	32.77	41.59	39.09	33.82	41.40
16	20.11	29.42	25.61	23.62	28.13	24.35
17	7.95	7.89	9.17	12.56	9.61	10.59
18	7.73	8.36	7.50	5.78	9.49	5.32
19+	21.22	21.56	16.13	18.96	18.94	18.35
Observations	2,238	19,966	2,078	2,406	20,850	2,149

Note: The Aberdeen sample is the subset of the "final sample" (defined in the text and Appendix A Table 1) with assignment scores in the interval [350,650). The LFS 01/02 (1) sample consists of all individuals born between 1950 and 1955 in the last 2 quarters of 2001 and the first 2 quarters of 2002. The LFS 2001/02 (2) sample restricts the LFS 2001/02 (1) sample to all individuals born in Scotland. LFS data is weighted using sampling weights. Number of weighted observations shown.

Appendix B Table 2: Aberdeen Study versus Labour Force Survey: highest educational qualification

	MEN			WOMEN		
	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland	Aberdeen	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland
None	21.54	21.08	22.71	24.85	26.99	31.50
Other qual.	2.86	9.92	7.14	3.49	9.38	5.41
Low CSEs	0.54	1.56	1.09	4.41	5.74	2.50
O level or equivalent	21.40	23.17	20.42	30.17	23.15	18.85
A level or equivalent	14.75	8.50	11.84	10.89	8.31	13.83
HNC, teaching, etc.	20.02	16.80	20.84	12.93	13.15	15.03
Degree	18.90	18.96	15.97	13.26	13.27	12.87
Observations	2,238	19,888	2,069	2,406	20,792	2,134

Note: see notes to Appendix B Table 1.

Appendix B Table 3: Labour Force Survey - percent with trade apprenticeship

	MEN		WOMEN	
	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland	LFS 01/02 (1) All UK	LFS 01/02 (2) Born in Scotland
None	26.17	35.06	5.02	6.86
Other qual.	0.00	0.00	0.00	0.00
Low CSEs	15.00	32.22	2.81	9.14
O level or equivalent	46.06	57.73	6.04	6.85
A level or equivalent	20.18	12.51	3.21	5.13
HNC, teaching, etc.	62.89	71.31	9.16	7.00
Degree	10.21	7.24	3.34	3.16
Total	30.66	37.68	4.84	5.85
Observations	19,790	2,061	20,737	2,131

Note: In the LFS trade apprenticeships are recorded in a separate question and are not included among the qualifications listed above. Numbers show the percentage of individuals with the corresponding level of qualification who also hold a trade apprenticeship. LFS samples as described in Appendix B Table 1.

Appendix B Table 4: Aberdeen Study versus Labour Force Survey: returns to education - gross annual income (Aberdeen), gross weekly pay (LFS)

	Aberdeen		LFS 01/02 (1) All UK		LFS 01/02 (2) Born in Scotland	
			MEN			
Years of post-compulsory education	0.072		0.069		0.085	
	(0.004)		(0.003)		(0.010)	
Left ft education at 15 or earlier		-				
Left ft education at 16		0.200		0.144		0.053
		(0.027)		(0.020)		(0.064)
Left ft education at 17		0.266		0.308		0.325
		(0.039)		(0.028)		(0.081)
Left ft education at 18		0.413		0.313		0.332
		(0.041)		(0.032)		(0.099)
Left ft education at 19+		0.458		0.539		0.579
		(0.028)		(0.023)		(0.073)
Observations	2152	2152	4,142	4,142	407	407
			WOMEN			
Years of post-compulsory education	0.128		0.110		0.129	
	(0.006)		(0.005)		(0.013)	
Left ft education at 15 or earlier		-				
Left ft education at 16		0.186		0.165		0.086
		(0.035)		(0.027)		(0.086)
Left ft education at 17		0.326		0.293		0.198
		(0.041)		(0.035)		(0.094)
Left ft education at 18		0.368		0.419		0.271
		(0.056)		(0.037)		(0.123)
Left ft education at 19+		0.801		0.755		0.816
		(0.036)		(0.031)		(0.084)
Observations	2316	2316	4,595	4,595	459	459

Note: Cells show least squares estimates of the returns to years of post-compulsory education. Dependent variable is log gross annual income for the Aberdeen sample and the log of the gross weekly wage for the LFS samples. Samples restricted to employees in work at the time of the survey. The Aberdeen sample consists of the entire sample of individuals replying to the 2001 survey. Aberdeen and LFS samples as described in Appendix B Table 1, now further restricted to employees in work at the time of the survey.

Appendix C: A Model of School Quality with Vocational Training

In this Appendix we present a simple model to support the argument made in section X, that for men, the existence of vocational training likely explains the absence of elite school effects on income. We begin with a baseline model without vocational training. This is adapted from the model that Card and Krueger (1996) used to examine the labor market implications of attending different school systems (our focus is on different types of school within the same system). We then introduce vocational training into the model.

C1: Baseline Model without Vocational Training

Modifying the Card and Krueger (1996) model slightly, we assume that individuals that have reached the compulsory school leaving age choose between leaving school and continuing in academic education for a further A years. We assume that for individual i that attended school type $s \in \{Nonelite, Elite\}$, this choice is made to maximize the following utility function:[1]

$$
\begin{aligned}
U(y_{is}, A_{is}) &= lny_{is} - f(A_{is}) \\
lny_{is} &= \theta_i + \theta_s + b_s^A A_{is} + u_{is} \\
f(A_{is}) &= \gamma_s^A c_i A_{is} + \frac{k}{2} A_{is}^2
\end{aligned}
$$

where y_{is} is annual earnings, θ_i is person-specific ability and c_i is the person-specific cost of academic education. We make the standard assumption that

[1] Card and Krueger (1996) assume that:

$$
\begin{aligned}
U(y_{is}, E_{is}) &= lny_{is} - f(E_{is}) \\
lny_{is} &= a_i + b_s E_{is} + u_{is} \\
f(E_{is}) &= c_i E_{is} + \frac{k}{2} E_{is}^2
\end{aligned}
$$

where E is total years of education (including compulsory school years).

$Cov(\theta_i, c_i) < 0$. The remaining parameters capture the effect of school type on the productivity of the (compulsory) years spent in school (θ_s), the return to additional years of academic education (b_s^A) and the cost of additional years of academic education ($\gamma_s^A c_i$).

It seems reasonable to allow the return to additional schooling to depend on the type of school attended: as Card and Krueger (1996) noted, a high-quality education may improve a student's ability to benefit from additional education. There are two reasons why it seems reasonable to allow the costs of additional education to depend on school type. First, since some of the post-compulsory education that took place in our setting occurred within the elite schools (i.e., students from non-elite schools had to transfer in), this may have created additional costs for non-elite students. Second, more generally and more plausibly, while the majority of elite school students stayed in academic education, the majority of non-elite school students did not, such that it might have been less costly for elite students to comply with default behavior than for non-elite school students to defy it (e.g., because of the costs of being separated from friends).

Maximization reveals the optimal schooling choice to be $A_i^* = max\{\frac{b_s^A - \gamma_s^A c_i}{k}, 0\}$ and maximized utility to be:

$$U(y_{is}, A_{is}^*) = \theta_i + \theta_s + max\{\frac{(b_s^A - \gamma_s^A c_i)^2}{2k}, 0\} + u_{is}$$

Proposition 1 summarizes three implications of this model.

Proposition 1

Assume the following conditions hold:

C1: The returns to academic education are higher for students that attended an elite school ($b_E^A > b_N^A$).

C2: The cost of academic education is lower for students that attended an elite school ($\gamma_E^A < \gamma_N^A$).

In that case:

1. There is some cost cutoff below which all individuals will pursue some academic education and above which no individuals will pursue any academic education. Among the students that pursue academic education, the length of academic education is decreasing in cost.

2. Elite school students will pursue more post-compulsory education.

3. Elite school students will obtain higher wages.

Proof

The first claim follows from inspection of the expression for A^*. The second follows from this expression and the assumption that $E[c_i]$ is the same for elite and non-elite students (among the borderline students). The third follows from substituting this expression into the equation for wages.

C.2: Vocational Training

We introduce vocational training by allowing students to choose between two post-secondary tracks: academic and vocational. Conditional on choosing the vocational track, we assume students solve a maximization problem similar to the one presented above, but with parameters b_s^A and γ_s^A replaced with parameters b_s^V and γ_s^V. We make the following assumptions on these parameters:

A1: $b_s^A > b_s^V$

A2: $\frac{b_s^A}{\gamma_s^A} < \frac{b_s^V}{\gamma_s^V}$

The first will ensure that the lowest-cost individuals (in expectation the most-able individuals) will choose academic training. The second will ensure that students on the margin of choosing vocational training over leaving school without pursuing any education will prefer vocational training to academic education.[2]

Proposition 2

1. Given assumptions A1 and A2, schooling decisions can be characterized by two cutoffs c_L and c_M. Students with $c_i < c_L$ will pursue academic education, with the length of academic education decreasing in cost; students with $c_L < c_i < c_M$ will pursue vocational training, with the length of vocational training decreasing in cost; students with $c_i > c_M$ will leave school without pursuing any vocational training or academic education.

2. An increase in b_s^A or a decrease in γ_s^A will increase the fraction of students that pursue academic education and decrease the fraction that pursue vocational training, with the fraction that leave school without pursuing any vocational training or academic education unchanged.

3. An increase in b_s^V or a decrease in γ_s^V will increase the fraction of students that pursue vocational training and decrease the fraction that leave school without pursuing any vocational training or academic education.

Proof

The proposition can be proved with reference to Figure 1. In particular, we

[2]It seems plausible to suppose the return to vocational training is lower than the return to academic education, since vocational training can be thought of as a combination of education and unskilled work. It seems plausible to suppose that the cost of vocational training is lower than the cost of academic education since vocational training pays a training wage.

can show that $c_M = \frac{b_s^V}{\gamma_s^V}$ (i.e., the type indifferent between vocational training and leaving school) and we know that $U(V_i^*; c_M) = 0$ while $U(A_i^*; \frac{b_s^A}{\gamma_s^A}) = 0$, where $c_M > \frac{b_s^A}{\gamma_s^A}$. We know that $U(A_i^*; 0) = \frac{(b_s^A)^2}{2k} > U(V_i^*; 0)$ and we can show that $U(A_i^*; c_i)$ and $U(V_i^*; c_i)$ cross at most once over the range $c_i \in [0, c_M]$.[3] The second and third parts of the Proposition then follow from Figure 1.

Proposition 3

If, in addition to conditions C1 and C2 and assumptions A1 and A2, we have the following condition:

C3: The costs of vocational training are higher for elite-school than non-elite school students

then:

1. Students assigned to elite school will pursue more post-compulsory academic education
2. Students assigned to elite school will pursue less vocational training
3. Assignment to an elite school need not increase wages

Proof

The first two claims follow immediately from Figure 2. The expected wage return to attending an elite school can be expressed as follows, where $\triangle_i \equiv lny_{Ei} - lny_{Ni}$:

[3] Otherwise, since the difference between them is continuous, and since it is positive when $c_i = 0$ and negative when $c_i = c_M$, there would be have to be two turning points. The first-order condition for a turning point demonstrates that there can be at most one value of c_i in this range.

$$
\begin{aligned}
E(\Delta_i) &= \int_0^{c_{L(N)}} [(b_E^A)^2 - \gamma_E^A b_E^A c_i - (b_N^A)^2 + \gamma_N^A b_N^A c_i] f(c_i) dc_i \\
&+ \int_{c_{L(N)}}^{c_{L(E)}} [(b_E^A)^2 - \gamma_E^A b_E^A c_i - (b^V)^2 + \gamma_N^V b^V c_i] f(c_i) dc_i \\
&+ \int_{c_{L(E)}}^{c_{M(E)}} [(b^V)^2 - \gamma_E^V b^V c_i - (b^V)^2 + \gamma_N^V b^V c_i] f(c_i) dc_i \\
&+ \int_{c_{M(E)}}^{c_{M(N)}} [-(b^V)^2 + \gamma_N^V b^V c_i] f(c_i) dc_i
\end{aligned}
$$

It is straightforward to construct examples in which the net effect is negative.[4] The intuition is that assignment to an elite school has ambiguous effects on human capital, increasing it for some (lower-cost) students that would anyway be inclined to academic study and decreasing it for other (higher-cost) students that would have pursued vocational training had they been assigned to the non-elite school.

C.3: Measured returns to education

An obvious question is whether the model can account for any of the other facts presented. We show that it can account for the lower return to academic education measured for men. To see why, note that:

$$
\begin{aligned}
E[lny_i|A_{is}] &= E(\theta_i|A_{is}) + \theta_s + b_s^A A_{is} + b^V E(V_{is}|A_{is}) \\
&= E(\theta_i|A_{is}) + \theta_s + b_s^A A_{is} + b^V (cons + r_{AV} A_{is}) \\
&= E(\theta_i|A_{is}) + \theta_s + (b_s^A + r_{VA} b^V) A_{is}
\end{aligned}
$$

[4] To construct an example in which the net effect is negative, suppose $c_i \sim U[0, \overline{c}]$ where $\overline{c} > c_{M(N)}$, such that:

$$
\begin{aligned}
E(\triangle_i) &= \frac{1}{k\overline{c}} \{[(b_E^A)^2 - (b_N^A)^2] c_{L(N)} + [(b_E^A)^2 - (b^V)^2](c_{L(E)} - c_{L(N)}) - (b^V)^2 (c_{M(N)} - c_{M(E)}) \\
&+ [\gamma_N^A b_N^A - \gamma_N^V b^V] \frac{c_{L(N)}^2}{2} + [\gamma_E^V b^V - \gamma_E^A b_E^A] \frac{c_{L(E)}^2}{2} - \gamma_E^V b^V \frac{c_{M(E)}^2}{2} + \gamma_N^V b^V \frac{c_{M(N)}^2}{2} \}
\end{aligned}
$$

If $b^V = 0.08$, $b_N^A = 0.1$, $b_E^A = 0.12$, $\gamma_N^V = 0.28$, $\gamma_E^V = 0.35$, $\gamma_N^A = 1.4$, $\gamma_E^A = 1.3$, $\overline{c} = 0.3$, $k = 0.01$ and $\theta_E = \theta_N$, then it is simple to show that $E(\triangle_i) \sim -0.01$.

$$= \quad E(\theta_i|A_{is}) + \theta_s + [b_s^A - \frac{E(V_{is}|A_{is}=0)}{E(A_{is}|A_{is}>0)} b^V] A_{is}$$

The presence of vocational training has two effects on the estimated returns to education. First, it generates the bias represented by the second term in square brackets. It can be seen that this will be zero if $b^V = 0$ (since $E(V_{is}|A_{is} = 0) = 0$), but positive otherwise. If c is distributed uniformly, then:

$$\begin{aligned} E(V_{is}|A_{is}=0) &= [\frac{b_s^V - \gamma_s^V(\frac{c_L+c_M}{2})}{k}] \frac{c_M - c_L}{\bar{c} - c_L} \\ E(A_{is}|A_{is}>0) &= [\frac{b_s^A - \gamma_s^A(\frac{c_L}{2})}{k}] \\ Bias &= b^V [\frac{b_s^V - \gamma_s^V(\frac{c_L+c_M}{2})}{b_s^A - \gamma_s^A(\frac{c_L}{2})}] \frac{c_M - c_L}{\bar{c} - c_L} \end{aligned}$$

Using the same parameters described above, it can be shown that this bias is on the order of 25 percent of the true return to academic education.

Second, vocational training weakens the ability bias generated by the first term $E(\theta_i|A_{is})$. Intuitively, that is because vocational training weakens the correlation between costs (hence ability) and academic education. Both forces imply that the measured returns to academic education will be smaller in the presence of vocational training.