ECON-C4100 - Capstone: Econometrics I

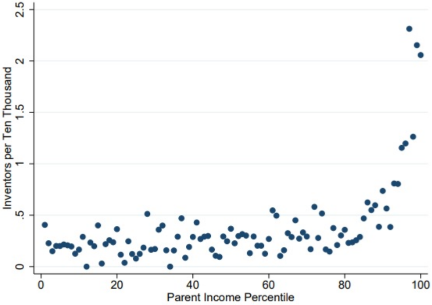
Lecture 11B: Aghion, Akcigit, Hyytinen & Toivanen: Parental education and invention

Otto Toivanen

Parental education and invention

Figure: Parental income and Prob(invent)

1A. 1930s U.S.

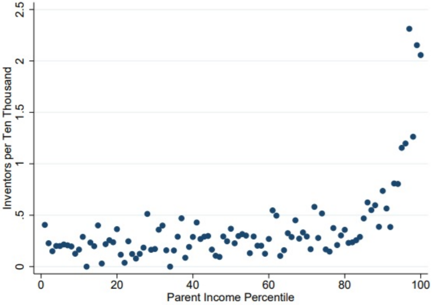


Note: Akcigit, U., Grigsby, J. & Nicholas, T. (2017). The rise of american ingenuity: Innovation and inventors of the golden age [National Bureau of Economic Research WP 23047].

Parental income and invention

Figure: Parental income and Prob(invent)

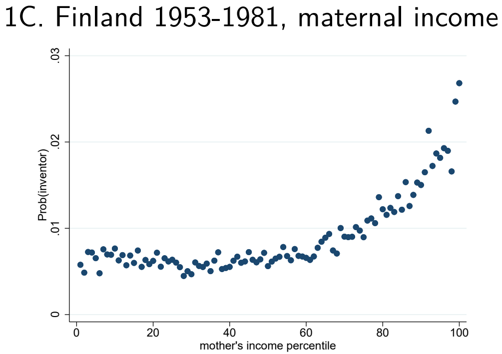
1B. 1980s U.S



Note: Bell, A., Chetty, R., Jaravel, X., Petkova, N. & Van Reenen, J. (2019). Who becomes an inventor in america? the importance of exposure to innovation. *Quarterly Journal of Economics*, *134*(2), 647–713

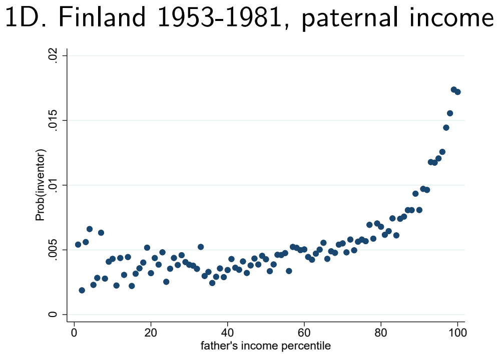
Parental income and invention

Figure: Parental income and Prob(invent)



Parental income and invention

Figure: Parental income and Prob(invent)



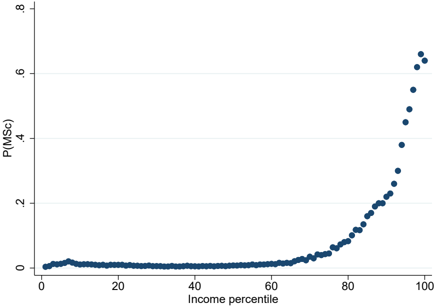
Finnish enigma

- How come in Finland the relationship between parental income and probability of offspring becoming an inventor is so similar to the US?

Parental income and education

Figure: Parental income and parental education

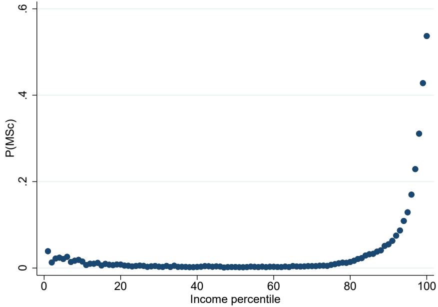
2A. Finland 1953-1981, maternal income & education



Parental income and education

Figure: Parental income and parental education

2B. Finland 1953-1981, paternal income & education



What do AAHT do?

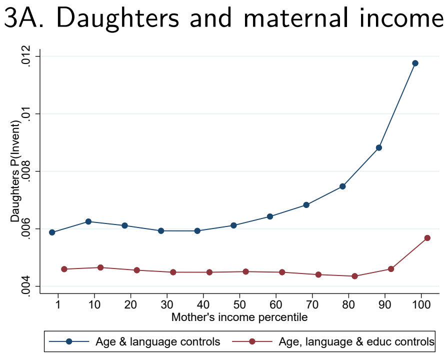
- How does the relationship between parental income and probability of becoming inventor change when parental education is controlled for?
- IV regression of probability of becoming inventor on parental education.

OLS regression

$$y_i = \boldsymbol{X_i'}\beta + f(income_{p,i}, \boldsymbol{\theta}) + g(Educ_{p,i}, \gamma) + \epsilon_i \tag{1}$$

- y_i is a dummy for being an inventor
- $\boldsymbol{X_i'}\beta$ are control variables and the associated vector of parameters to be estimated
- $f(income_{p,i}, \boldsymbol{\theta})$ is a fifth order polynomial of income of the parent of type p ($p = mother,\ father$), with $\boldsymbol{\theta}$ being the associated vector of parameters to be estimated
- $g(\boldsymbol{Educ_{p,i}}, \gamma)$ includes a vector of field (STEM, non-STEM) and level (secondary, college, masters, PhD level, with base-level being omitted) of education dummies $Educ_{p,i}$ of parent of type p, with γ being the associated vector of parameters to be estimated
- ϵ_i is the error term

Parental income and education



IV

- Instrument: Parental distance to nearest university from birth-municipality, measured in the year when the parent in question turns 19
- Exclusion restriction: parental distance to university uncorrelated with unobservables affecting probability of offspring becoming an inventor.

IV

Our main estimation equation is of the form

$$y_i = \boldsymbol{X_i'}\beta + \delta D_i + \epsilon_i \tag{2}$$

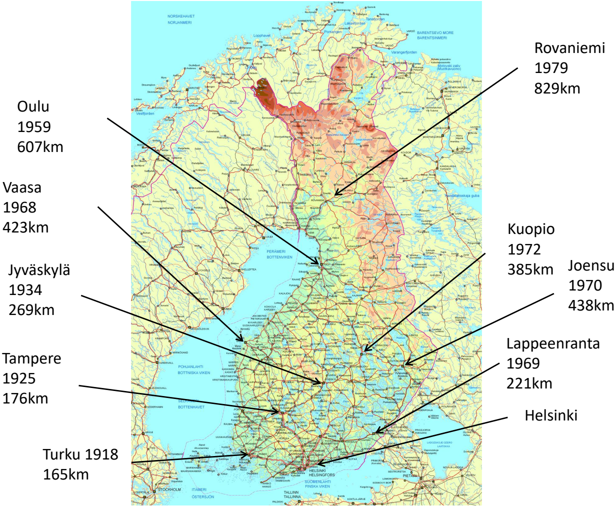
- y_i is the outcome dummy variable taking value 1 if individual i is an inventor of a patent, and 0 otherwise
- $\boldsymbol{X_i}$ is a vector of controls (maternal and paternal year of birth dummies, a dummy for mother tongue not being Finnish, and the controls for the birth municipalities of both parents discussed above); β is the associated coefficient vector
- D_i is the parental education dummy taking value 1 if individual i has at least one parent with at least an MSc and 0 otherwise
- δ is the causal parameter of interest and
- ϵ_i is an error term capturing all those determinants of an individual becoming an inventor that are unobservable to us

Challenge with IV

- Parents growing up near a university are different from those growing up further away.
- Solution #1: utilize data around the establishment of new universities
- Solution #2: bring in control variables that reduce/remove the potential problem.

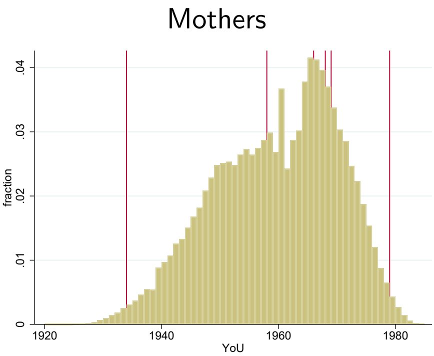
Parental income and education

Figure: Map of Finnish university establishments 1918 - 1979



Parental income and education

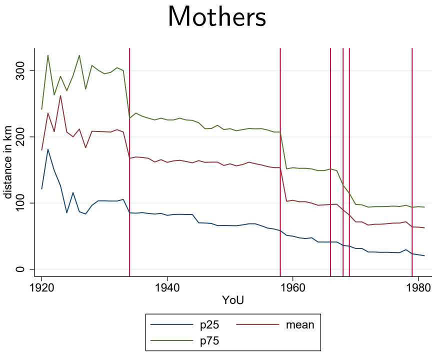
Figure: Distribution of parents by year at age 19



Note: *YoU* = year of university (age 19)

Parental income and education

Figure: Distribution of parents by year at age 19



Note: *YoU* = year of university (age 19)

Parental income and education

Table: Distance correlations

Parent	$P(inventor)$	D(MSc parents)	MSc_p	Count	MSc_{cohort}
Maternal	-0.0110	-0.0360	0.0179	0.1088	-0.1958
	(0.1679)	(0.0000)	(0.0251)	(0.0000)	(0.0000)
Paternal	-0.0221	-0.0135	-0.0117	0.0766	-0.1548
	(0.0078)	(0.1039)	(0.1590)	(0.0000)	(0.0000)

Parent	p50	p90	IQ
Maternal	-0.2042	-0.1395	-0.0452
	(0.0000)	(0.0000)	(0.0028)
Paternal	-0.2336	-0.1227	-0.0536
	(0.0000)	(0.0000)	(0.0007)

Note: reported numbers correlation coefficient and p-value. All other variables pertain to parent, or parental muni-year cohort, but IQ is the son's IQ

Parental income and education

Table: Estimation results

	Panel A. All Children			
	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
D(MSc parents)	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	251.04	497.453	108.49
Nobs		1 450 789		
	Panel B. Daughters			
D(MSc parents)	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	251.04	497.453	108.49
Nobs		709 117		
	Panel C. Sons			
D(MSc parents)	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	251.04	497.453	108.49
Nobs		741 671		
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Conclusions

- Parental education has a positive causal impact on probability of offspring becoming inventors
- Effect larger in absolute terms for sons, in relative terms for daughters
- Results survive when using IQ as additional control
- Effect larger for cohorts just before than for cohorts just after comprehensive school reform
- Results robust in a number of ways: different samples, different outcome variables, different measures of parental education, different functional forms...
- The fact that estimated coefficient varies as the instrument is changed suggests that we identify a **Local Average Treatment Effect**, or LATE