Public Economics
Education

Iñigo Iturbe-Ormaetxe

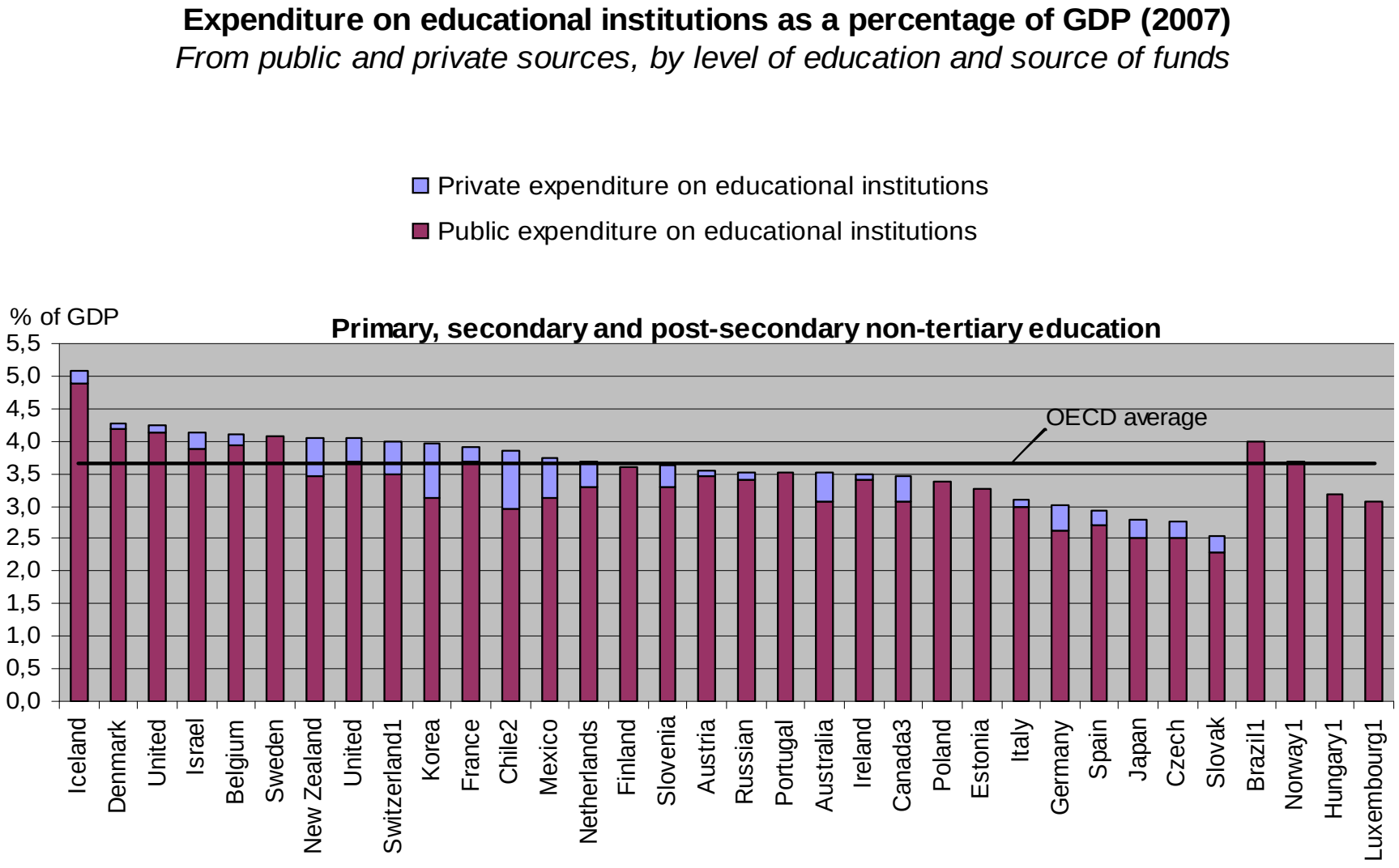
U. of Alicante

2011-12

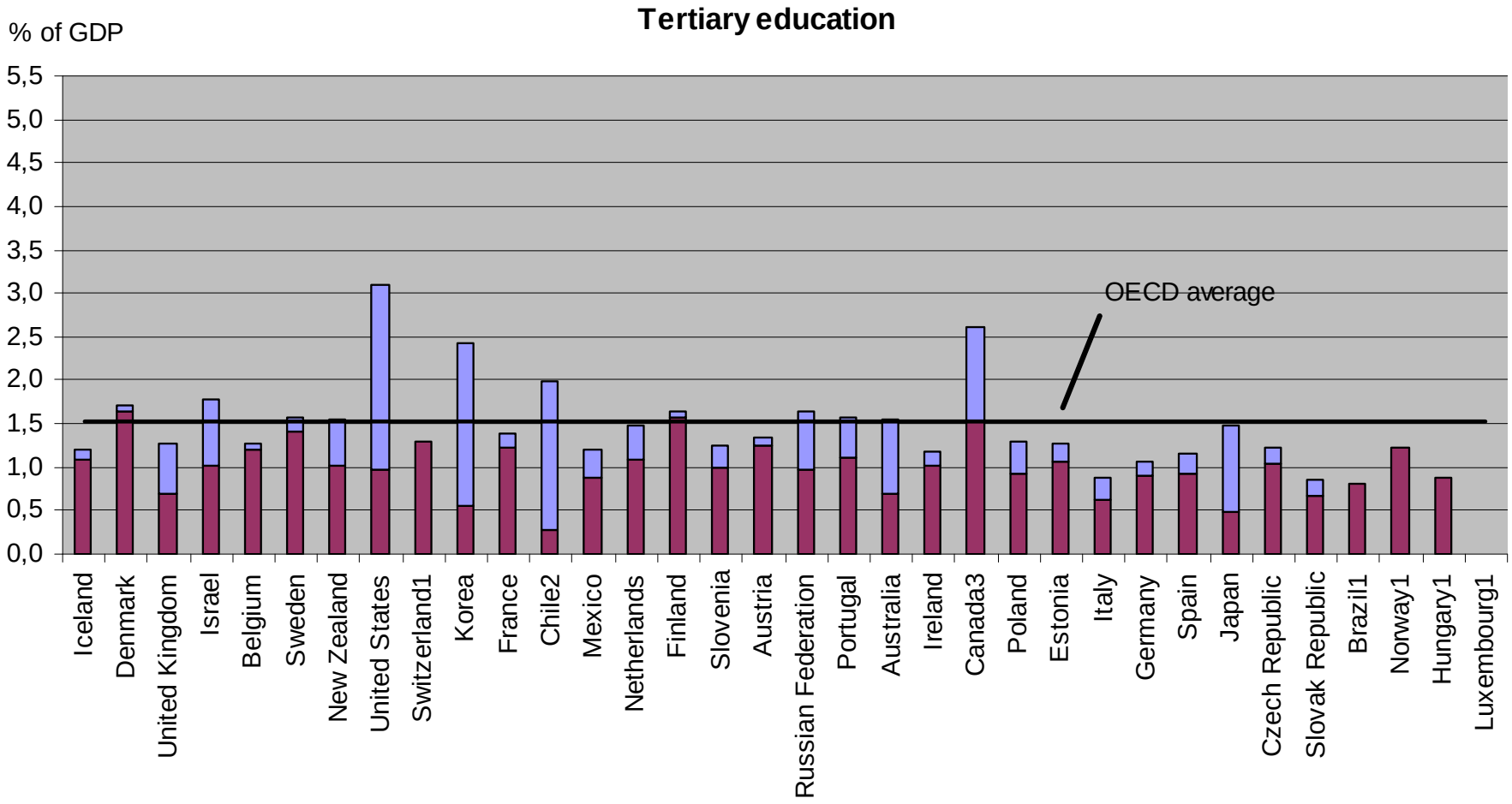
Introduction

- Education is one of the major public programs. All governments across the world are concerned about education
- Average expenditure in the OECD countries is around 5.5 % of GDP
- Slightly below 4 % in primary, secondary and post-secondary education
- Around 1.5 % in tertiary education

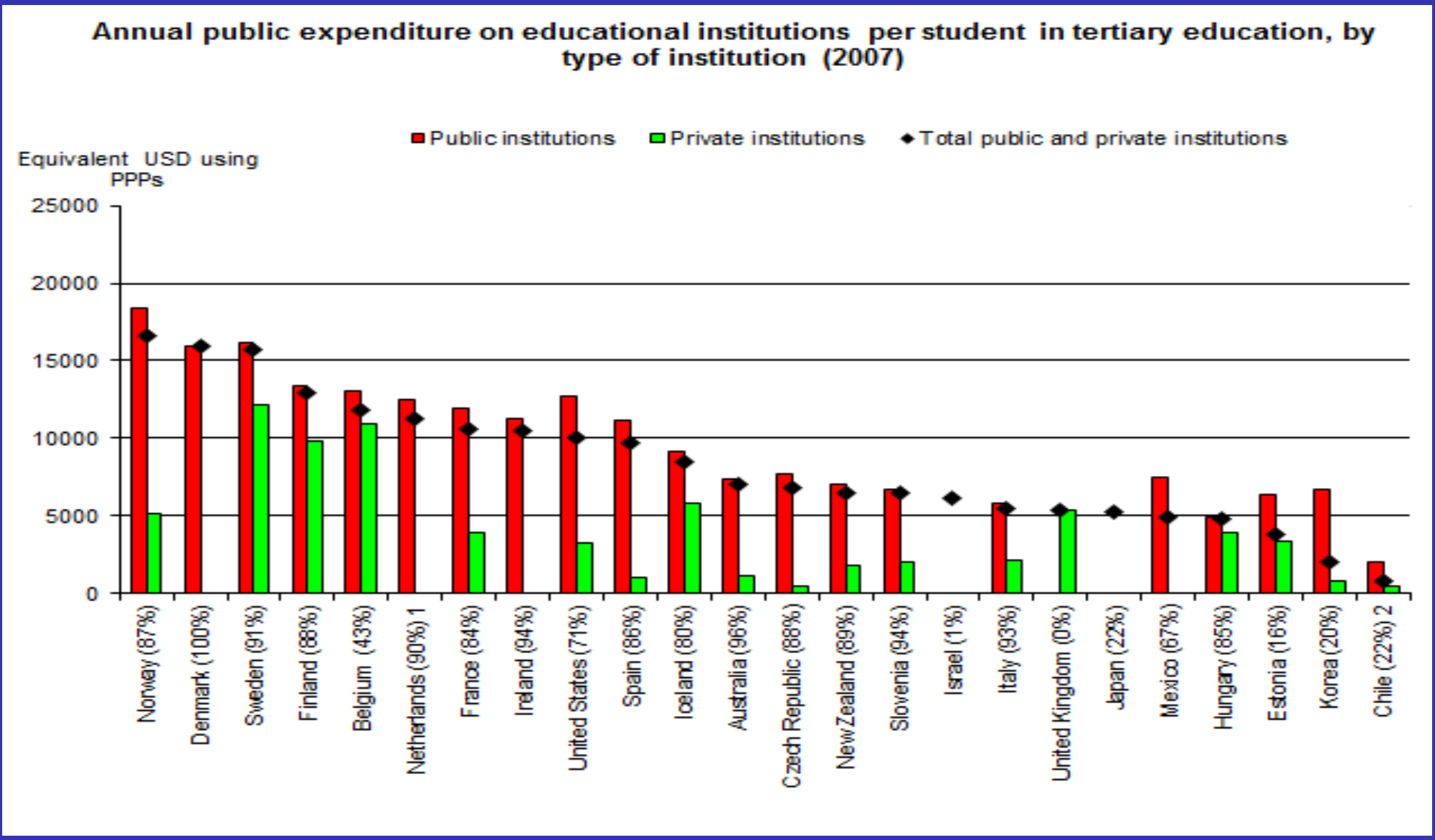
Expenditure before tertiary education



Expenditure in tertiary education



Annual expenditure, tertiary education



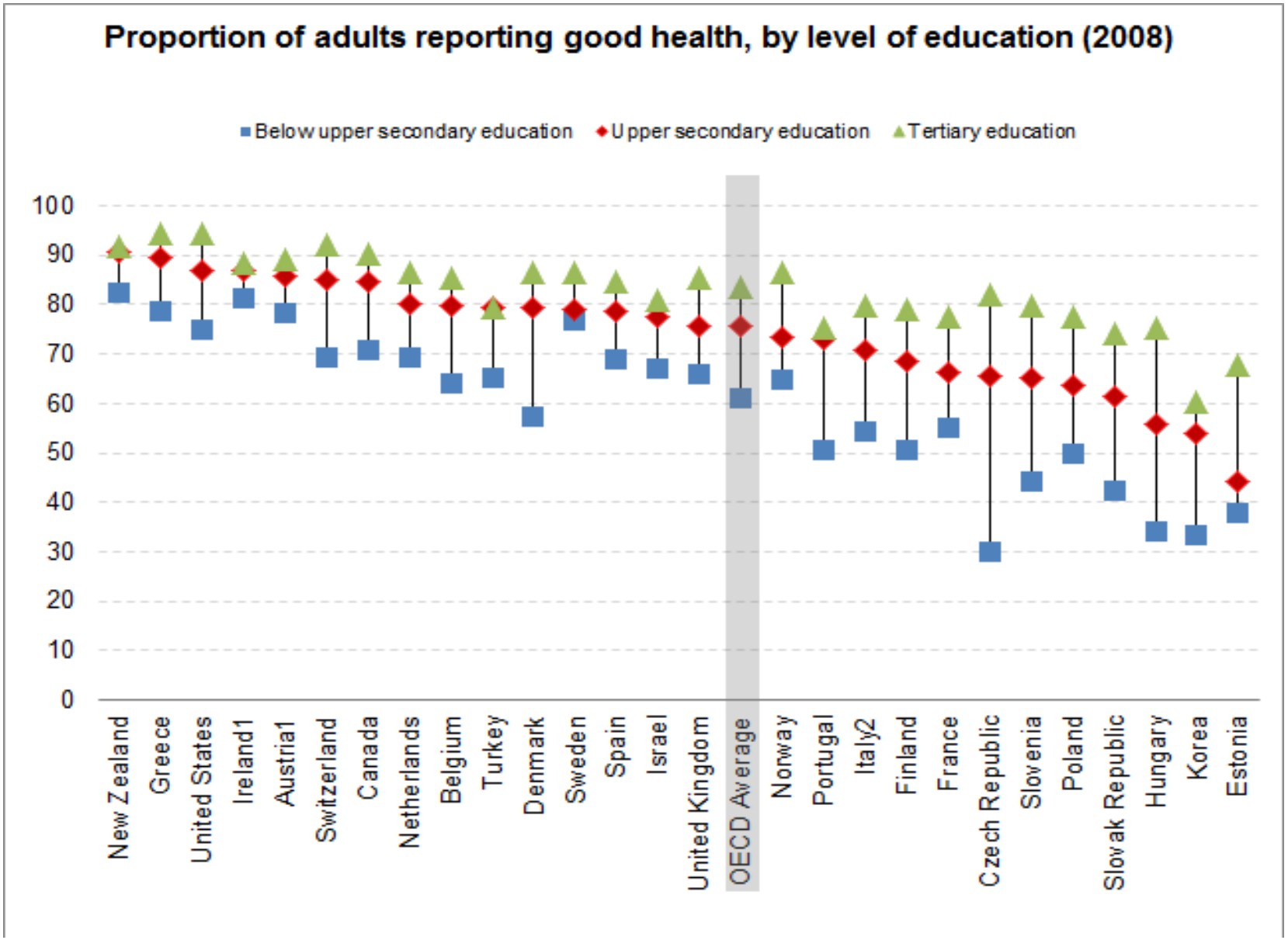
Not a pure public good

- Education is not a pure public good
- Remember: pure PG defined by non-excludability and non-rivalry
- There are congestion effects and the marginal cost of providing education to an additional child is far from zero.
- The marginal and average costs are approximately the same for large numbers of students. Also, there is no technical difficulty in charging individuals
- Then, why should intervene governments?

Why government intervention?

- One justification is the existence and importance of **externalities**
- Education seems to generate large externalities. It has been shown to:
  - Reduce crime
  - Improve health
  - lower mortality
  - increase political participation
- Increasing in schooling seems to reduce most types of adult crime (Lance Lochner, 2011). This means a reduction of negative externalities

Health by level of education



Why government intervention?, cont.

- However, there is also a very large private return to being able to read, so without government support, many individuals would learn this and other basic skills
- The question is: suppose there is no intervention and individuals choose privately the amount of education they wish. Now the government intervenes and individuals choose more education. Would this intervention generate any significant externalities?
- Lochner and Moretti (2004) estimate that high school completion may lower annual social costs of crime by $3,000 per male graduate. Increasing high school completion by 1 % would save more than $2 billion
- They also calculate benefits from reduction in mortality between $1,500-2,500 per additional graduate

Borrowing constraints

- There is an indirect externality to investments in **science and technology** education: people with these scarce skills are the key to technological progress; typically innovators capture only a fraction of their overall contribution to the increase in productivity
- A reason why people do not invest as much in education as they would like is the existence of **borrowing constraints**
- Another reason that applies for primary and secondary level is that parents are who take decisions on behalf of their children
- Although most parents are willing to spend money in their children's education and behave altruistically, others may not. It is not obvious that all parents take the best decisions for their children
- If only private education exists, the children of such parents might receive an insufficient education. This gives a rationale for public support of **primary and secondary** education: children's access to education should not depend on their parents' financial ability or their sense of altruism

Provision vs. financing

- The above arguments provide a rationale for government financing of education, **but not for government production**
- In the US government production dominates at the primary and secondary level, but not at college level
- In Europe, it also dominates at the college level
- In Spain, private education ("concertada") accounts for 1/3 of primary and secondary education

Decision to acquire Human Capital

- Assume earnings with s years of post-compulsory schooling is $W(s)$
- No direct costs of education, only foregone earnings
- Individuals live forever
- The present discounted value of s years of education is:

$$PDV = \int_s^{\infty} e^{-rt} W(s) dt = \frac{1}{r} W(s) e^{-rS}$$

Education decision

- Taking logs we can write this expression as:

$$\log(PDV) = \log W(s) - rs - \log r$$

- The first-order condition is:

$$\frac{W'(s)}{W(s)} = r$$

- This requires to acquire education up to the point where the increase in log earnings is equal to the rate at which future earnings are discounted

Equilibrium

- If all individuals are identical and there are different levels of education in equilibrium, it must be the case that, for any pair $s \neq s'$:

$$\log W(s) - rs - \log r = \log W(s') - rs' - \log r$$

- In particular, this is true for $s' = 0$. Then:

$$\log W(s) = \log W(0) + rs$$

- If we include years of education on the right-hand side of the earnings function, the coefficient of s is a measure of the return to education

Measuring the rate of return

- For the US a typical OLS estimate from an earnings function is about 8 %
- In some advanced countries may be as high as 17-20 %
- This suggests that education is a very good investment, since few other investments offer an 8 % real return
- This return is higher for more able people (Taber (2001)) and for children from better backgrounds (Altonji and Dunn (1996))
- Those from better backgrounds and higher ability are also more likely to attend college

Estimates of rates of return: Trostel, Walker and Woolley (2002)

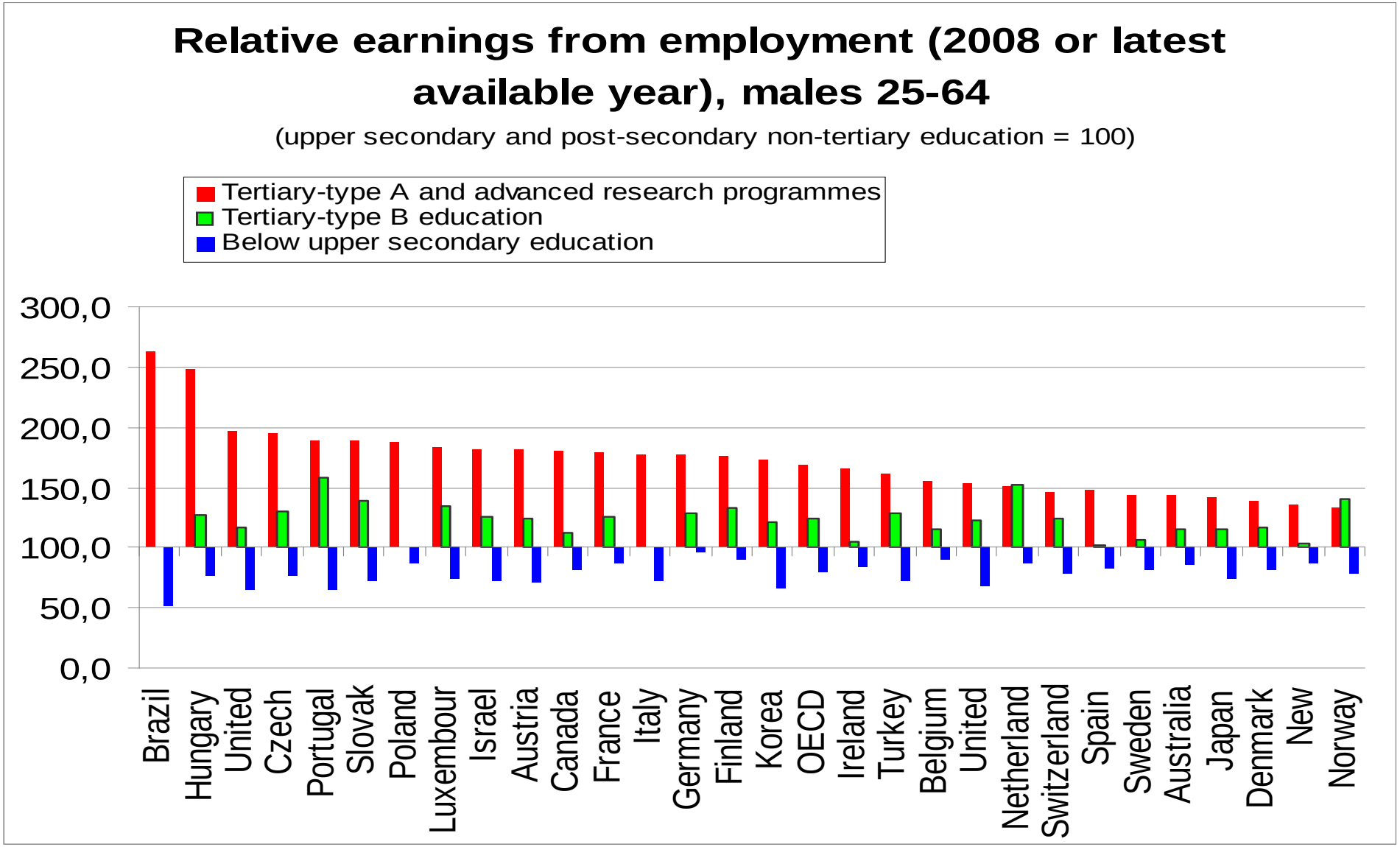
OLS Estimates of Returns to Years of Education

Country	Males		Females	
USA	0.074	*0.004*	0.096	*0.005*
Great Britain	0.127	*0.006*	0.130	*0.006*
Russia	0.044	*0.004*	0.053	*0.004*
Norway	0.023	*0.002*	0.025	*0.003*
Australia	0.051	*0.004*	0.052	*0.006*
Netherlands	0.031	*0.002*	0.019	*0.004*
Austria	0.038	*0.004*	0.064	*0.006*
Poland	0.073	*0.005*	0.100	*0.005*
New Zealand	0.033	*0.004*	0.029	*0.005*
Italy	0.037	*0.003*	0.053	*0.005*
Ireland	0.085	*0.006*	0.090	*0.008*
Japan	0.075	*0.007*	0.094	*0.014*
Hungary	0.075	*0.007*	0.077	*0.006*
Canada	0.038	*0.008*	0.045	*0.008*
Spain	0.046	*0.005*	0.038	*0.010*
Philippines	0.113	*0.015*	0.192	*0.030*
Pooled	0.048	*0.001*	0.057	*0.001*

Source: Trostel, Walker and Woolley (2001). International Social Survey data.

Notes: Robust standard errors are in italics. Include year dummies, union status, marital status, age and age squared and, in the case of the aggregate equation, country-year dummies.

OECD data


Relative earnings from employment (2008 or latest available year), males 25-64
(upper secondary and post-secondary non-tertiary education = 100)
Tertiary-type A and advanced research programmes
Tertiary-type B education
Below upper secondary education
300,0
250,0
200,0
150,0
100,0
50,0
0,0
Brazil
Hungary
United
Czech
Portugal
Slovak
Poland
Luxembour
Israel
Austria
Canada
France
Italy
Germany
Finland
Korea
OECD
Ireland
Turkey
Belgium
United
Netherland
Switzerland
Spain
Sweden
Australia
Japan
Denmark
New
Norway

Puzzle

- If the rate of return to education is so high, why people does not acquire more education?
- Possible answers are:
  - Liquidity constraints (cannot borrow against future HC)
  - OLS estimates are upward biases, and rates or return are not really that high
  - Heterogeneity in the returns to education
  - Investing in education entails risk

Bias in OLS estimate

- Why OLS estimate can be biased?
- Most common answer is that education does not generate higher income. Instead, individuals with higher ability choose to receive more education and receive higher wages. They are paid more because they have high ability. This is called **ability bias**
- Ability has an effect on wages independent of education, but is positively correlated with schooling. If this is not controlled for in the regression we get an estimate that is upward biased

Omitted variable bias

- Suppose the true model is:

$$\ln(W) = b_0 + bs + b_2 a + \varepsilon,$$

where a is ability

- But a is omitted from the regression and we estimate by OLS:

$$\ln(W) = \beta_0 + \beta s + u$$

- We get:

$$p\lim\left(\widehat{\beta}_{ols}\right) = b + \frac{Cov(a,s)}{Var(s)},$$

- If a and s are positively correlated, OLS estimate is biased upwards

Solutions to omitted ability bias

- Include better controls for ability in the regression. For instance, use data on IQ tests
- Twin studies
- Instrumental variables

Using IQ as a proxy for ability

- Using data for 935 men from the National Longitudinal Survey of Youth (NLSY), USA 1980:

$$\ln(W) = \underset{(0{,}112)}{5{,}503} + \underset{(0{,}007)}{0{,}078}s + \underset{(0{,}003)}{0{,}020} \times \text{Exp},$$

  where Exp is experience. We show standard errors
- Including a proxy variable for ability (IQ obtained in a test taken before finishing compulsory education):

$$\ln(W) = \underset{(0{,}122)}{5{,}198} + \underset{(0{,}007)}{0{,}057}s + \underset{0{,}003}{0{,}020} \times \text{Exp} + \underset{(0{,}001)}{0{,}006} \times \text{IQ}$$

- The estimated return to education falls from 7.8 % to 5.7 %

Twin studies

- Studies using brothers or twins are based on the fact that brothers or twins share at least one common component of the unobservable ability variable
- A model for earnings of twin 1 and twin 2 in pair i :

$$\begin{aligned} \ln(W_{1i}) &= b_0 + bs_{1i} + b_2 a_{1i} + \varepsilon_{1i} \\ \ln(W_{2i}) &= b_0 + bs_{2i} + b_2 a_{2i} + \varepsilon_{2i} \end{aligned}$$

- If s is correlated with ability, we cannot estimate by OLS

Twin studies

- But, if we assume that identical twins have $a_{1i} = a_{2i}$
- We take differences to get:

$$\Delta \ln(W_i) = b_0 + b\Delta s_i + \Delta\varepsilon_i$$

- Now the regressor is uncorrelated with the error term and the OLS estimate will be consistent
- Estimates using twins are typically lower

Problems with twin studies

- Small sample size
- Measurement error in schooling biases the coefficient of schooling toward zero
- First differences generally magnify the importance of measurement error bias. That is, increases the tendency for estimates of schooling to be attenuated (biased toward zero)

Brief econometric detour

- Suppose the model to estimate is:

$$y = x^* \beta + \varepsilon,$$

  but instead of observing x^* we observe $x = x^* + u_x$, where u_x is purely random
- Then, we can write:

$$y = x\beta + (\varepsilon - \beta u_x)$$

- The presence of u_x generates a mechanical correlation between the error term $(\varepsilon - \beta u_x)$ and the explanatory variable $(x = x^* + u_x)$

Brief econometric detour

- It is easy to prove that:

$$
\begin{aligned}
p\lim(\widehat{\beta}) &= \frac{cov(y,x)}{var(x)} = \frac{cov(x^*\beta+\varepsilon, x^*+u_x)}{var(x^*+u_x)} \\
&= \frac{var(x^*)}{var(x^*+u_x)}\beta + \frac{cov(x^*\beta, u_x) + cov(\varepsilon, x^*+u_x)}{var(x^*+u_x)} \\
&= \frac{var(x^*)}{var(x^*)+var(u_x)}\beta
\end{aligned}
$$

- Now $\widehat{\beta}$ converges in probability to a fraction $\frac{var(x^*)}{var(x^*)+var(u_x)} < 1$ of the true β. This is called attenuation bias since $\widehat{\beta}$ is biased toward zero

First differences

- If the variables above are replaced with their first differences:

$$p\lim(\widehat{\beta}^{FD}) = \frac{var(\Delta x^*)}{var(\Delta x^*) + var(\Delta u_x)}\beta$$

- The variance of the noise term doubles from $var(u_x)$ to $var(\Delta u_x) = 2var(u_x)$. However, the variance of the signal falls if there is substantial serial correlation in x^*
- If the autocorrelation in x^* is ρ_x, then:

$$var(\Delta x^*) = 2(1 - \rho_x)var(x^*).$$

This is smaller than $var(x^*)$ if $\rho_x \geq 1/2$. That is, unless $\rho_x \leq 0$, very unlikely, first differences will tend to magnify the attenuation bias due to measurement error

Solution

- Use a second measure of the variable that is measured with error
- Provided that measurement error in the two measures is uncorrelated, the second measure can be an instrument for the first measure
- Ashenfelter and Krueger (1994) asked each twin about the other twin's schooling and use this variable as an instrument
- That is, they instrument ΔS with the difference in the cross-reported level of S, $\Delta S'$, assuming classical measurement error, $\Delta S' = \Delta S + \Delta v$
- If $\Delta S'$ is a valid instrument, $\beta_{WTIV} = \beta$

Monozygotic twins (Bingley, Christensen, Walker 2009)

Table 1 *Recent MZ Twins Estimates in the Literature*

Study	Sample	Date	Country	Gender	# twin pairs	β_{OLS}	B_{WT}	β_{WTIV}
Ashenfelter and Krueger (1994)	Twinsburg	1991	US	Pooled	147	0.084 (0.014)	0.092 (0.024)	0.129 (0.030)
Berhman *et al* (1994)	NAS-NRC	1973	US	Pooled	141	0.094[a] (0.011)	0.035 (0.004)	0.101 (0.012)
Miller *et al* (1995)	Australian Twins Register	1985	Australia	Pooled	602	0.064 (0.002)	0.025 (0.005)	0.048 (0.010)
Ashenfelter and Rouse (1997)	Twinsburg	1991-93	US	Pooled	333	0.110 (0.009)	0.070 (0.019)	0.088 (0.025)
Berhman and Rosenzweig (1997)	Minnesota Twins Register	1993	US	Pooled	720	0.113[a] (0.005)	0.104 (0.017)	n.a.
Miller, Mulvey and Martin (1997)	Australian Twins Register	1985	Australia	Male	282	0.071[d] (0.003)	0.023 (0.008)	0.033 (0.014)
				Female	320	0.057[d] (0.002)	0.028 (0.006)	0.058 (0.011)
Rouse (1998)	Twinsburg	1991-93, 95	US	Pooled	453	0.105 (0.008)	0.075 (0.017)	0.110 (0.023)
Isacsson (1999)	Swedish Twin Registry	1990	Sweden	Pooled	2492	0.046 (0.001)	0.022 (0.002)	0.024[b] (0.008)
Isacsson (2004)	Swedish Twin Registry	1990	Sweden	Pooled	2609	0.066[c] (0.009)	0.028[c] (0.009)	0.052[c] (0.036)
Bonjour *et al* (2004)	St Thomas' Hospital twins register.	1999	UK	Female	187	0.077 (0.001)	0.039 (0.023)	0.077 (0.033)

Notes: Table 1 from Bound and Solon (1999) and Table 6 from Card (1999) updated. a – GLS estimate. b – not instrumented but evaluated at a reliability ratio of 0.88. c– evaluated at upper secondary level of schooling. d – pooled DZ and MZ.

IV approach

- Another solution is to instrument years of schooling
- Suppose our model is:

$$\ln(W) = \beta_0 + \beta s + u$$

- A suitable IV z must meet 2 conditions:
  1. Relevance: the instrument has to be correlated with years of schooling ($cov(z, s) \neq 0$)
  2. Exogeneity: the instrument has to affect wage only through the channel of schooling, and therefore must be uncorrelated with the error term in the wage equation ($cov(z, u) = 0$)
- Note that (1) can be tested, but (2) cannot

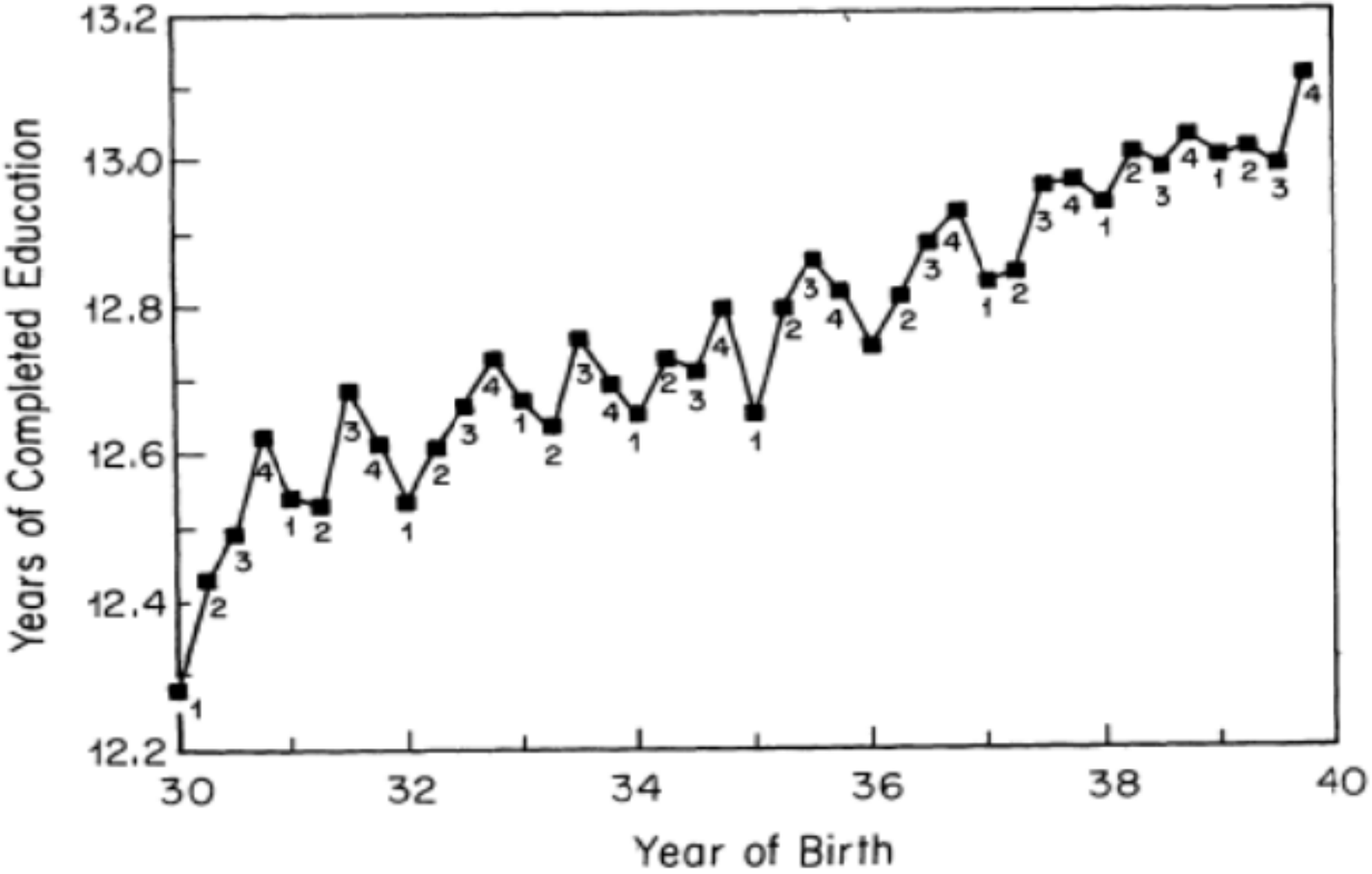
Candidates

- The last digit of the social security number (fails (1))
- IQ (a measure of): fails (2) since it is highly correlated with omitted ability
- Mother's education has been used in several papers. It is positively correlated with child's education (it satisfies (1)), but it can be also correlated with child's ability (it may fail (2))
- Number of siblings while growing up. More siblings is associated with lower average levels of education

Angrist and Krueger (1991)

- Angrist and Krueger (1991) came up with a binary instrumental variable for schooling using census data on men in the US
- The IV variable z is a dummy variable that takes value 1 if the man was born in the first quarter of the year, and zero otherwise
- It seems that quarter of birth should be uncorrelated with the error term ($cov(z, u) = 0$). But it is also needed that z has to be correlated with schooling ($cov(z, s) \neq 0$)
- This seems to be the case. Why? Because of compulsory school attendance laws

Quarter of Birth and Schooling



Years of Education and Season of Birth
1980 Census

School attendance laws

- Since compulsory schooling laws are based on age, the age at which students begin school depends on quarter of birth
- In many states, students start school in the calendar year in which they turn 6
- A student born in January will generally start 1st grade nearly a year younger than a student born in December. January-born students tend to be in 9th grade when they reach the end of compulsory schooling and can choose to drop out, while December-born students tend to be in 10th grade when they can drop out
- Quarter of birth is an instrument for schooling as long as we assume that any difference in wages of those born in different quarters is a result only of differences in schooling
- Angrist and Krueger also find that for students who finish high school, there is no relationship between years of schooling and quarter of birth

IV estimations

- Since the correlation between years of schooling and quarter of birth is small, they need a very large sample size to get precise IV estimates
- They use a sample of 247,199 men born between 1920 and 1929. The OLS estimate of the return to education is .0801 (standard error .0004), and the IV estimate is .0715 (standard error .0219)
- Note the t statistic for the OLS estimate (about 200) is much larger than the t statistic for the IV estimate (3.26). The IV estimate is statistically different from zero, but its confidence interval is much wider than the one of the OLS estimate

Angrist and Krueger: Criticisms

- Bound, Jaeger, and Baker (1995) defend that quarter of birth can be correlated with unobserved factors that affect wage
- Some studies have found that quarter of birth is correlated with (among others):
  - School attendance rates
  - Behavioral difficulties
  - Performance in reading, writing, math
  - Individuals born early in the year are more likely to suffer from schizophrenia
  - There is evidence of variation by quarter of birth in the incidence of mental retardation, autism, dyslexia...
- This implies that quarter of birth may have a direct effect on wages. Even a small amount of correlation between z and u can produce serious problems for the IV estimator

Another brief econometric detour

- Suppose that z and u are correlated. Then:

$$p\lim(\widehat{\beta}^{IV}) = \beta + \frac{corr(z,u)}{corr(z,x)}\frac{\sigma_u}{\sigma_x},$$

  where we refer to population moments. Even if $corr(z,u)$ is small, the inconsistency of the IV estimator can be large if $corr(z,x)$ is also small ("weak instruments" problem)
- Comparing with the OLS estimator:

$$p\lim(\widehat{\beta}^{OLS}) = \beta + corr(x,u)\frac{\sigma_u}{\sigma_x}.$$

- It is better to use the OLS estimator when $\frac{corr(z,u)}{corr(z,x)} < corr(x,u)$

Card (1995)

- Card (1995) uses as instrument a dummy variable ("nearc4") that takes value 1 if the individual grew up near a four-year college
- He assumes it is uncorrelated with the error term (reasonable?), and finds that it is correlated with schooling
- He runs a regression in which schooling is the endogenous variable and has nearc4 among other regressors. He gets:

$$\widehat{s} = \underset{(0{,}24)}{16{,}64} + \underset{(,088)}{,320} \times \text{nearc4} + ..$$

- The coefficient of nearc4 says that those who lived near a college in 1966 had, on average, one-third of a year more of education than those who did not grow up near a college

Card (1995)

- Card estimates the effect of schooling on wages both by OLS and IV with the above instrument
- He gets $\widehat{\beta}^{OLS} = ,075$ (standard error ,003) and $\widehat{\beta}^{IV} = ,132$ (standard error ,055)
- The 95 % confidence interval for the IV estimate is $[,024, ,239]$, which is very wide. This is the price to pay in order to get a consistent estimator

Other issues on returns to education

- Until now we have focused on quantity of schooling, but quality and type of schooling is also important
- Below we explore the effect of class size
- Also interesting to study why schooling raises earnings
- There are two main explanations:
  - Human capital view (Becker)
  - Signaling view (Spence)

Human capital

- According to the HC view, education increases skills, and then wages
- Human capital refers to the **intangible stock of skills that are embodied in people.** These skills can be innate (like talent) or can be acquired through education, training and learning
- In the second case, the acquisition of human capital is a form of investment
- The bottom line is that **human capital should be viewed as any other form of capital.** To acquire it has a cost, and it provides a stream of future returns in the form of higher earnings and welfare
- Individuals at school do not only acquire knowledge, but also some behavioral aspects like punctuality, reliability, perseverance, etc. that are valuable in the workplace (see Weiss (1995))

Screening and signalling

- According to the signalling view, one of the important functions of education is to **identify** the abilities of different individuals
- Individuals who go to school longer get a higher wage, but not because schools have increased their productivity. It is because the schools have identified them as being more productive
- The school is a **screening device**, separating the able and motivated from the less able and less motivated
- Also students are interested in **signalling** their ability to employers

Screening and signalling, cont.

- Social returns to education are far less than private returns. In the extreme case (where schooling does not affect productivity at all), and it serves only to identify who is more able and who is less able, the social return from education would be zero, while private returns are not zero
- However, even if the only effect of schooling would be to screen workers, there can be social returns, since education enables a better matching between individuals and jobs

HC vs. Signalling

- There is controversy between the human capital view and the screening view
- There is some empirical evidence that persons with more schooling earn more even when there is no opportunity to signal
- For example, more educated farmers are more productive, as the self-employed
- Moreover, returns at macro level are of the same magnitude as conventional micro estimates
- Finally, there is a literature on employer learning. Fabian Lange (2007) estimates that employers learn quickly. He obtains an upper bound on the contribution of signaling to the gains from schooling of around 25 %

Other controversies in education

- In the US and other countries there is a debate on whether increased educational expenditures leads or not to increased education performance (Pisa Report)
- US is among the countries with the highest levels of expenditures per pupil. However, this does not translate into good results, as measured by international test scores
- One of the issues is the effect of class size. Some studies support the view that smaller class size has a positive impact on student performance: Krueger (1999) using data from the Tennessee STAR experiment (see the book by Angrist and Pischke, Mostly Harmless.., 2009)
- Angrist and Lavy (1999) use quasi-experimental evidence from Israel and find a strong link between class size and achievement
- The fact that others find no evidence of this link is because class size is endogenous to students' type. Below in Lazear's (2001) paper

Other controversies, cont.

- Some authors claim that the real outcome of education is not higher scores, but higher productivity and higher wages
- But even if schooling has little impact on earnings, because family is critical, it is possible to redistribute resources towards the disadvantaged, to offset these differences in background
- The problem is, how to allocate resources. Is it better to direct more to the top, since those at the top are the most productive, or is it better to direct more to the bottom to shorten the wage gap and reduce inequalities?
- Also, should we invest more when the person is young or in later ages? (Heckman below)

Vouchers

- Another debate concerns **vouchers:** each child would be given a coupon (a check) to be used at the school of the parent's choice
- Under this proposal, public schools would have to compete directly with private schools
- This is a way of separating 2 things. Today the government **both** finances education and produces it
- Some authors believe that government is not an efficient producer of education, private schools would do much better. Also, parents become more committed to that school and more involved in their children's education
- This improves school performance. Advocates of vouchers argue that with vouchers not only the rich have the opportunity to choose, but also the poor

Vouchers, critics

- Not all parents are well informed (but this will always happen)
- In some areas there is limited choice of schools
- Private schools are not better. A selection effect works. Those who send their children to private schools are more committed to education. This commitment accounts for the differences in performance
- The voucher system can lead to a more stratified society, with children of wealthy and well-educated parents going to private schools and children of poor and less educated parents going to public schools (this is what happens now in Spain, without vouchers!)

Higher education

- The fundamental difference between higher education and primary and secondary education is that students have already reached an age at which they can make decisions for themselves
- In this respect, the role of the government should be to ensure access. This is done by many channels
- First, most governments subsidize higher education, charging in public institutions tuitions that are a small fraction of total costs (0-20 % of total costs)
- Second, most governments provide grants to cover tuition costs, other costs and even foregone earnings
- Third, governments sometimes also sponsor loan programs and provide tuition tax deductions and credits

Higher education, cont.

- Subsidies have been criticized as being untargeted
- Since enrollment rates are typically higher among the children of the rich, on average it is the richer people who benefit from such subsidies (although progressivity of income tax may reduce this effect)
- Moreover, as the subsidies can be thought of as directed at the children themselves, lifetime incomes, even of children from poor families, will be higher than the average income of the population
- Some authors say that government should focus on loans or grants to children of poor families

Class size

- Empirical evidence on the effect of class size on students's achievement is not conclusive. Lazear (2001) proposes a simple explanation
- Classroom teaching is a public good and congestion effects are important
- A student who is disruptive or takes up teacher time affects, not only his learning, but that of the others in the class. He creates a **negative externality** on his classmates
- Class size is a choice variable (an endogenous variable) and **the optimal class size varies inversely with the attention span of the students**
- It is efficient to use fewer teachers and a higher student-teacher ratio when students are better behaved

Model

- We call p the probability that any given student **IS NOT** impeding his own or other's learning at any moment of time. It can be seen as the probability that the student is behaving properly
- If class size is n, the probability that all students are behaving is p^n, and disruption occurs $1 - p^n$ of the time
- Clearly, p must be high. For example, if $p = ,98$ and $n = 25$, $1 - p^n = 0{,}4$. Disruption happens 40 % of the time
- If $n = 20$, then $1 - p^n = 0{,}33$
- Note: $\frac{\partial(1-p^n)}{\partial n} = -p^n \ln(p) > 0$ and $\frac{\partial^2(1-p^n)}{\partial n^2} = -p^n \left[\ln(p)\right]^2 < 0$

Model, cont.

- Now we ask how much a student would pay to be in a class of size n
- The value of one unit of learning for each student is V, determined by the market value of HC
- A school of Z students with m teachers and m classes has to choose class size. Every student gets out of the class $Vp^{\frac{Z}{m}} = Vp^n$. Class size is $n = \frac{Z}{m}$
- The cost of a teacher is W. A private school wants to maximize profits. Profits are:

$$\pi = ZVp^{\frac{Z}{m}} - Wm$$

- Revenue is the value of total human capital produced

First-order condition

- Dividing by Z, profit per student is:

$$\frac{\pi}{Z} = Vp^n - \frac{W}{n}$$

- The foc is:

$$-V\frac{Z^2}{m^2}p^{Z/m}\ln(p) - W \leq 0$$

- Or simply $Vp^n \ln(p) + W/n^2 = 0$. Free entry of firms in the market for education results in zero profits at the equilibrium in the long-run

Second-order condition

- We also compute the second derivative of profits:

$$VZ^2 p^{Z/m} \ln(p) \frac{2m + Z\ln(p)}{m^4}$$

- Since $\ln(p) < 1$, we need $2m + Z\ln(p) > 0$. This will be the case if p is close to 1

Comparative statics

- Using the foc we can derive some comparative statics:
- Optimal class size rises with teacher's wage and falls with V
- The most important result is that optimal class size rises with the probability that students behave well. It is optimal to reduce class size when students are less well-behaved
- An example to illustrate the effect of p. We must price W relative to V
- In equilibrium, W relative to V must be sufficiently low. Otherwise, private schools could not exist

Optimal class size

- Using profits per student, we obtain an upper bound for W
- Profits per student are $Vp^n - \frac{W}{n}$. This is positive if:

$$W \leq nVp^n$$

- Suppose that the ratio W to Vp^n (Vp^n is what each student gets out of the class) is 5
- If $Z = 100$ and $p = ,99$, we get from the foc that $m = 3{,}94$, which implies $n^* \approx 25$. At the optimum, each student gets out an amount $Vp^{n^*} = ,99^{25} = 0{,}777821$
- If class size is $n = 27$, she gets out $,99^{27} = ,762342$. From 25 to 27, educational output per student reduces by only a 2 %

Optimal class size, cont.

- This illustrates why it is so difficult to find significant class effects
- What is more important, class size is a choice variable. Because class size is a choice variable, researchers often observe small, or possibly even positive class size effects
- As n^* declines with p, better-behaved students are in larger classes
- Although more disruptive students are in smaller classes, the effect of reducing class size is not sufficient to overcome their deficiencies
- This is Proposition 2 in the paper

Proposition 2

- After optimal class-size adjustment, educational output per student is higher in larger classes with better-behaved students than in smaller classes with less well-behaved students
- As an example, when $p = ,99$, $n^* = 25$ and educational output is .78
- When p falls to .98, $n^* = 19$, and educational output is .68
- The effect of the lower value of p is not offset by the reduced class size

Class size effects

- Since class size is inversely related with W, large class-size effects are most likely to be observed when the cost of teachers is low
- Low W implies low optimal class sizes
- Reducing class size has a larger effect on educational output in small classes than in larger ones
- Preschool teachers are less expensive than college professors, which generates small class size for preschoolers
- Class-size effects should be more important in preschool classes than they are at the college level

Low p students

- The effect of class-size reductions is greater for low p students
- Suppose that we reduce class-size from nk to n $(k > 1)$. The percentage increase in educational output from that reduction is:

$$\frac{\left(p^n - p^{nk}\right)}{p^{nk}} = p^{n-nk} - 1$$

- Differentiating with respect to p we get $(n - nk)\, p^{n-nk-1} < 0$
- Another implication is that very poorly behaved children (with sufficiently low p) cannot be accommodated by a private school. There is a value p^* such that $n^* = 0$ for $p < p^*$
- In the public sector, both schools and students might be forced to provide education, even for very low p students. If the social planner is concerned only with efficiency, she would set n as large as possible

Heterogeneous students

- Suppose there are 2 types of students: A's and B's and assume $p_A > p_B$. Is it better to segregate student by type or to mix them?
- First assume that all classes are of size n. The proportion of A's is α
- Output per student in a school with segregated classes is $\alpha p_A^n + (1-\alpha) p_B^n$
- With integrated schools, it is $p_A^{\alpha n} p_B^{(1-\alpha)n}$
- The difference is:

$$\alpha p_A^n + (1-\alpha) p_B^n - p_A^{\alpha n} p_B^{(1-\alpha)n}$$

Heterogeneous students, cont.

- When $p_A = p_B$, this difference is zero
- Differentiating the expression above with respect to p_A, we get:

$$\frac{\partial}{\partial p_A} = \alpha n p_A^{n-1} \left(1 - \frac{p_B^{(1-\alpha)n}}{p_A^{(1-\alpha)n}} \right) > 0,$$

  for $p_A > p_B$
- The result extends immediately to the case in which class size can differ

Policies to foster HC, by Heckman

- Learning is a lifetime affair and much learning takes place outside of schools
- Skill formation in early pre-school years is of vital importance. And in those years families play the most important role
- Families and environments play the crucial role in motivating and producing educational success:
  "Failed families produce low-ability, poorly motivated students who do not succeed in school. Policies directed toward families may be a more effective means for improving the performance of schools than direct expenditure on teacher salaries or computer equipment."
- Also post school learning is an important source of skill formation. It accounts for as much as one third to one half of all skill formation in a modern economy

Non-cognitive skills

- Heckman criticizes the preoccupation with achievement tests and measures of cognitive skills as indicators of success of an educational intervention
- There is a full array of socially and economically valuable non-cognitive skills and motivation produced at schools. For example, early intervention programs do not substantially alter I.Q., but they raise non-cognitive skills and social attachment of participants (e.g., crime rates)
- Another error is the belief that abilities are fixed at very early ages. Schooling produces ability and ability creates a demand for schooling

Parents

- There is a fundamental mistrust among policy makers of the wisdom of parents to choose wisely if offered choices about their children's education
- Also there is mistrust of competition and incentives as means of improving the performance of schools
- According to the available evidence, when offered choices and the opportunity to experiment, most parents generally choose wisely, at least after they have gained some experience:
  "They can distinguish the good teachers from the bad teachers, and the good schools from the bad schools"
- Heckman examines the merits of some of the recent policy proposals by examining the evidence-or lack of evidence-supporting them

Questions

- It seems to be very little evidence of unexploited externalities in Western countries. They could be important in a world in which there is no educational policy, but not in Western countries where current subsidy of direct costs to students at major public colleges is around 80 %. It is crucial to recognize the status quo, and consider policies as changes from that status quo
- Another missing piece in current debates is the **need to prioritize**. We have to decide how to allocate the budget into different programs. The particular policy recommendation of Heckman is: "Invest in the very young and improve basic learning and socialization skills; subsidize the old and the severely disadvantaged to attach them to the economy and the society at large"

Questions, cont.

- The return of a euro spent on the young is higher than the return of a euro spent at a later age. The reason is that early investments are harvested over a longer horizon than those made later in the life-cycle
- Also human capital has dynamic complementarity features. Learning begets learning (reinforcement process)
- For a given opportunity cost of funds, an optimal investment strategy is to invest relatively less in the old and relatively more in the young
- The central message of Heckman is that efficiency in public spending would be enhanced if HC investment were directed more toward the disadvantaged young, and less toward older, less-skilled persons

Evidence on credit constraints

- In most countries the poor have a much lower college participation rate than the rich. The traditional interpretation is that of short-term family credit constraints
- Based on this, policies that further subsidize higher education for low income families have been advocated
- It is true that if families had to rely on their own resources to finance all of their schooling costs, the level of educational attainment in society would decline
- An alternative (but not exclusive) interpretation is that long-run family and environmental factors play a decisive role in shaping the abilities and expectations of children. That is, the decision on college is not taken at 18 but much earlier

Credit constraints, cont.

- This influence accumulates over many years to produce ability and college readiness. If finances of the poor prevents them from providing decent primary and secondary schooling for their children, and produce a low level of college readiness, public policies aimed at reducing the short-term borrowing constraints for the college expenses is unlikely to be effective
- According to Heckman: *"Only to the extent that the family income of able high school graduates falls below levels required to pay for college will short-term credit constraints hinder college entry. Given the current college financial support agreements that are available to low income and minority children in the US, the phenomenon of bright students being denied access to college because of credit constraints is an empirically unimportant phenomenon"*

Primary and secondary schools

- All operating costs are completely subsidized, except the opportunity costs
- There is a debate on whether we should pay the teachers more, or decrease class size. The available empirical evidence indicates that, **within current ranges of expenditure** in most Western countries, inputs such as class size and spending per pupil have little, if any, effect on the future earnings of students
- In particular, Heckman claims that the evidence for the US is that spending may be too high, such that returns turn to be negative!!
- To justify additional spending on primary and secondary education, we would need to appeal to other social benefits not captured by earnings. However there is much debate about this point

Early childhood investments

- Note that the evidence does not say that school quality does not matter. Increasing it from very low levels, it matters greatly. But there is little evidence that marginal improvements from **current** levels of schooling quality are likely to be effective
- Programs targeted toward early stages are highly effective in reducing criminal activity, promoting social skills and integrating disadvantaged people into mainstream society
- This is also the case with programs designed to keep adolescents in school. They are effective because they increase schooling and improve employment opportunities
- To sum up: "The role of the family is crucial to the formation of learning skills, and government interventions at an early age that mend the harm done by dysfunctional families have proven to be highly effective"

García-Peñalosa and Wälde (1994)

- We study first the traditionally tax-subsidy system
- The system performs poorly, because there is a trade-off between efficiency and equity
- Next we compare the tax-transfer system with three alternatives: a pure loan scheme, a loan scheme with income-contingent repayments, and a graduate tax system
- Criteria used are: Pareto efficiency, ex ante equality of opportunity and ex post equality of lifetime incomes

Model

- Population size is N. Individuals live 3 periods and differ in parents' income n (density is $f(n)$)
- In period 1, individuals get compulsory education. At the start of period 2, they choose whether to work or to invest in higher education
- In the first case they work as unskilled workers in periods 2 and 3
- If they go to college, they study in period 2 and work as skilled workers in period 3
- Consumption takes place only in period 3

Model, cont.

- Fixed cost of education E. Borrowing not possible
- Without public intervention, only those with $n \geq E$ attend college
- Production is $Y = F(H, L)$, where H is skilled labor and L unskilled labor $(H + L = N)$. Assume $F(0, L) = F(H, 0) = 0$
- Markets are competitive, hence factors are paid their marginal products $w_H(H, L)$ and $w_L(H, L)$
- Benchmark case of perfect capital markets

Perfect capital markets

- Individuals can borrow and there is no need for government intervention
- If an individual decides not to study, lifetime income is $W_L = (1+R)w_L(H,L)$, where R is the discount rate
- If she decides to study, it is $W_H = -E + Rw_H(H,L)$
- The equilibrium level of skilled workers is H^*, for which lifetime incomes are equalized. It is defined by:

$$(1+R)w_L(H^*, N-H^*) = -E + Rw_H(H^*, N-H^*)$$

- We call H^* the **efficient** level of human capital. It is efficient; there is ex post equality of lifetime incomes; there is ex ante equality of opportunity. In fact, all individuals are indifferent between going and not going to college

Traditional tax-subsidy systems

- Again individuals cannot borrow. All individuals pay T (a LST) in period 1. Those who study get a subsidy sE
- The budget constraint of the government is:

$$T = \frac{sEH}{N}$$

- Those who do not study pay a tax and get no subsidy. Those who study pay the tax and get the subsidy. This implies a net transfer to those who study, since $sE - T = sE(1 - \frac{H}{N}) > 0$
- Define $\underline{n}(s)$ as the minimum income required to be able to afford education. This is $\underline{n}(s) = (1 - s)E + T$. Substituting,

$$\underline{n}(s) = E\left(1 - s\left(\frac{N - H}{N}\right)\right)$$

Tax-subsidy system II

- If $s = 0$, we have $\underline{n}(s) = E$ as before
- If $s = 1$, then:

$$\underline{n}(s = 1) = E\left(1 - \left(\frac{N - H}{N}\right)\right) = E\frac{H}{N}$$

- Instead of $n \geq E$, now the requirement is $n - T \geq (1 - s)E$ or $n \geq (1 - s)E + T$
- Clearly $\underline{n}'(s) < 0$, the higher is s the more individuals can afford the cost
- We assume that under the subsidy s, the capital market constraint is still binding for some households. Then, lifetime income of graduate students exceeds lifetime income of unskilled workers

Tax-subsidy system III

- More individuals want to study than can afford to do so. The number of students is:

$$H = \sum_{n \geq \underline{n}(s)} f(n)$$

- Equations above determine jointly H and $\underline{n}(s)$ as a function of s and $f(n)$
- Suppose that without intervention (when $s = 0$) the level of human capital is $H_0 < H^*$. If government wants to implement H^*, it will fix a subsidy rate s^* such that there are exactly H^* individuals with income $n \geq (1 - s^*)E + T$. We call s^* the efficient subsidy

Efficient subsidy

- Since the reduction in the education cost due to the subsidy is greater than the tax paid, lifetime income of those who study increases, for a given H, and W_H shifts upwards
- At H^*, the difference between lifetime incomes is:

$$\begin{aligned} & W_H(s^*) - W_L(s^*) \\ = \; & -(1-s^*)E - T^* + Rw_H(s^*) - ((1+R)w_L(s^*) - T^*) \\ = \; & -(1-s^*)E + Rw_H(s^*) - (1+R)w_L(s^*) = s^*E \end{aligned}$$

- The efficient subsidy has two distributional consequences:
  1. A transfer of resources from poor to rich: all individuals pay the tax, but only the rich get the subsidy (**reverse redistribution**)
  2. Those who study enjoy a larger income than those who do not. The poor are systematically excluded

Equality of lifetime income

- If government wants to implement a policy that guarantees equality of lifetime income, the stock of human capital will be $\widehat{H}$, and the subsidy $\widehat{s}$, where:

$$(1+R)w_L(\widehat{H}, N-\widehat{H}) - \widehat{T} = -(1-\widehat{s})E - \widehat{T} + Rw_H(\widehat{H}, N-\widehat{H})$$

- It is clear that $\widehat{s} > s^*$ and $\widehat{H} > H^*$. The increase in the number of educated workers implies that their marginal product is lower than the social cost of education E plus the foregone wage. The level of H is inefficient. There are "too many" skilled workers
- The tax-subsidy system is characterized by an equity efficiency trade-off. The efficient subsidy implies inequality in lifetime incomes, and the equitable subsidy implies an inefficient amount of skilled labor. Besides, there is no equality of opportunities with the equitable subsidy

Distributional effects

- We want to know if a particular individual is better or worse off with the subsidy
- The tax-subsidy scheme has three effects: (1) A tax effect, as those who do not study have to pay a net tax which subsidizes students. (2) A cost-reduction effect, since those who engage in education pay less, even taking into account their tax payments. (3) A wage effect, since the increase of skilled workers results in a higher unskilled wage
- Consider the efficient subsidy. Middle-income individuals are always better off under the subsidy. Without intervention they could not afford education, and with it they can. Their lifetime wage is greater than without the subsidy (they could not study)

Distributional effects, cont.

- The richest individuals may be better off or worse off. Without the subsidy they could afford education. Now, the skilled wage is lower but the cost of education is reduced by the subsidy. The low-income agents can be better off or worse off. They pay taxes but the wage of unskilled workers is higher
- With an equitable subsidy, total output is lower than the efficient level. It can be even lower than before the subsidy. It could be that even the middle-income group is worse off

Adding uncertainty

- We look for better mechanisms of financing education
- Without uncertainty, the solution would be simple. Abolish subsidies completely and introduce a government loan scheme. We would attain all our objectives
- However, we have to take into account the uncertainty related to investments in human capital. Higher education is a risky investment
- We model uncertainty as follows. Suppose students have to take an exam at the end of their education. If they pass, they become skilled workers. If they do not pass, they work as unskilled workers. We assume $p \in (0, 1)$ is the probability of passing the test
- Also we assume wages are constant for both types of workers. Individuals are risk-averse and their preferences are represented by $U(\cdot)$ with decreasing absolute risk aversion

Who educates?

- If H is the number of students, there will be pH skilled workers in the next period
- We assume that the expected return to education net of the education cost is higher than the discounted wages from working as an unskilled worker for two periods, that is:

$$-E + R[pw_H + (1-p)w_L] > (1+R)w_L$$

- Under this condition the optimal value of H^* is N (all individuals should educate)

Pure loans

- Any individual who studies pays the full cost E, irrespective of whether or not she succeeds in education
- A worker who does not invest in education has net lifetime income $n + (1 + R)w_L$; if she invests but fails has $n - E + Rw_L$; if she invests and succeeds has $n - E + Rw_H$
- Expected net return is:

$$\begin{aligned} & p(n - E + Rw_H) + (1 - p)(n - E + Rw_L) \\ = \quad & n - E + R[pw_H + (1 - p)w_L] \end{aligned}$$

- We define as $G(n, s = 0)$ the **difference** between the expected utility of investing in education and not investing when there are no subsidies to education

Pure loans II

- It is defined as:

$$\begin{aligned} G(n,0) &\equiv (1-p)U(n-E+Rw_L) \\ &\quad +pU(n-E+Rw_H)-U(n+(1+R)w_L) \end{aligned}$$

- If $G(n,0)>0$ for some n, then $G(n',0)>0$ for all $n'>n$. This is an immediate consequence of declining absolute risk aversion
- Defining $\underline{n}^l$ as the level such that $G(\underline{n}^l,0)=0$, the total number of students will be:

$$H^l=\sum_{n\geq \underline{n}^l} f(n)$$

Pure loans III

- As long as $\underline{n}^l > 0$, $H^l < H^* = N$ which is the optimal one. So with loans, the allocation is not efficient
- Besides, it is not equitable **ex post**. The lifetime income of a skilled worker $(-E + Rw_H)$ is greater than that of an individual who never studied $((1+R)w_L)$ which is also greater than that of a student who did not succeed $(-E + Rw_L)$
- Finally, there is no equality of opportunity

A graduate tax

- The public sector offers loans and pre-finances part of the education through public debt. Total debt is repaid by levying a tax on those who finished college successfully
- Government subsidizes a fraction s of the cost of education EH and finances it by borrowing. Total tax revenue is TpH. The constraint is $sEH = RTpH$
- An individual who invests in education and succeeds, pays education cost of $(1-s)E + RT$ over her lifetime. Using the budget constraint:

$$(1-s)E + RT = \left(1 + s\frac{(1-p)}{p}\right)E > E$$

- Those who study but fail, pay only $(1-s)E$

A graduate tax, cont.

- Compared to the pure loan scheme, education cost raises for successful students and falls for unsuccessful ones. Expected net return from studying is:

$$\begin{aligned} & p(n-(1-s)E-RT+Rw_H)+(1-p)(n-(1-s)E+Rw_L) \\ = \; & n-(1-s)E-pRT+R(pw_H+(1-p)w_L) \\ = \; & n-E+R(pw_H+(1-p)w_L) \end{aligned}$$

- Exactly the same than with the pure loan system
- For an individual who studies, now the variance of lifetime income (the difference between success and failure) is lower. It is:

$$R(w_H-w_L)-RT=R(w_H-w_L)-\frac{sE}{p},$$

and with the loan system it was $R(w_H-w_L)$

A graduate tax, efficiency

- The difference between the expected utility of investing and not investing in education $G(n,s)$ is:

$$\begin{aligned} G(n,s) \equiv & (1-p)U(n-(1-s)E+Rw_L)+ \\ & +pU(n-\left(1+s\frac{(1-p)}{p}\right)E+Rw_H) \\ & -U(n+(1+R)w_L) \end{aligned}$$

- If $s=0$, we are back to the case of a pure loan scheme. This new system implies that the expected return of education is independent of s but its variance decreases with s. This implies that $G(n,s)$ will be increasing in s

A graduate tax, assessment

- Defining $\underline{n}^g$ as the level such that $G(\underline{n}^g, s) = 0$, we have that the total number of students will be:

$$H^g = \sum_{n \geq \underline{n}^g} f(n)$$

  As s increases, $\underline{n}^g$ moves to the left. However, as long as $\underline{n}^g > 0$, the optimum will not be attained
- However, it does not imply reverse distribution and it reduces differences between ex post lifetime incomes of the three types of workers compared to the pure loan scheme. There remain differences because parental wealth provides insurance in the event of unsuccessful studies
- Equality of opportunity is not attained

Income contingent loans (ICL)

- The student receives a loan from the state such that: (i) repayment only takes place if her income after graduation exceeds a pre-specified level, (ii) annual repayments do not constitute more than a certain proportion of her income, and (iii) repayments cease once the loan plus interests has been repaid
- Those who study borrow E. Those who succeed and pass the test must repay E
- A tax is raised to repay the educational costs of those who failed, $(1-p)HE$. Total revenue in present value is RTN (all individuals pay the tax). The constraint implies $RT = \frac{(1-p)HE}{N} < E$

Income contingent loans II

- Again we have:

$$\begin{aligned} G(n,0) &\equiv (1-p)U(n-RT+Rw_L)+pU(n-RT-E+Rw_H) \\ &\quad -U(n-RT+(1+R)w_L). \end{aligned}$$

- Expected net lifetime income from studying is:

$$\begin{aligned} & p(n-RT-E+Rw_H)+(1-p)(n-RT+Rw_L) \\ = \; & n-pE-RT+R[pw_H+(1-p)w_L] \end{aligned}$$

Income contingent loans, cont.

- This is greater than the expected net lifetime income under the graduate tax system. To see this, note:

$$-pE - RT + R[pw_H + (1-p)w_L] > -E + R[pw_H + (1-p)w_L],$$

since $-pE - RT > -E$, because $RT = \frac{(1-p)HE}{N} < (1-p)E$, and substituting we reach to $(1-p)H < (1-p)N$

- This system will be more efficient, since as the expected net lifetime income is higher it is less attractive not to study

ICL, lifetime income

- Difference in lifetime income between those who succeed and those who fail is $R(w_H - w_L) - E$
- Under the graduate income tax, it was $R(w_H - w_L) - \frac{s}{p}E$. The first will be higher when $s > p$. This makes this scheme less efficient
- Recall: it is better in expected return, but it is worse in terms of variance
- It is better in expected return at the cost of those who do not study and pay the tax RT. Successful students earn as much as under the pure loan scheme and more than under the graduate tax scheme. Unsuccessful students earn more than under the pure loan scheme and more or less than under the graduate tax scheme. They earn less with the income contingent loan if $RT > (1-s)E$, or $s > \frac{N-(1-p)H}{N}$. This will hold for a high enough value of s

ICL vs. PL

- The income contingent loan also outperforms the pure loan scheme. Provided s is close to one, it is the more efficient. Also it avoids reverse redistribution
- What if the government subsidizes a fraction of the cost out of the general tax revenue and finances the remaining part with a system of income-contingent loans?

The size of the subsidy

- The graduate tax entails a moral hazard problem
- In general, p is not independent of the individual. The subsidy need not be so large that lifetime income after successfully completing education $(-\left(1+s\frac{(1-p)}{p}\right)E+Rw_H)$ is the same than failing $(-(1-s)E+Rw_L)$
- Both will be equal if the subsidy is s_f (full insurance) such that:

$$s_f\frac{E}{p}=R[w_H-w_L]$$

- Both will get the same expected income $-E+R[pw_H+(1-p)w_L]$. If $s^f>1$, a subsidy of $s=1$ would avoid the moral hazard problem. If not, the problem is more complicated