e-learning course

Encouraging the further development of Higher Education Act

Theme 3

Admission to HE should reward quality and promote high academic standards in pre-tertiary education: Quality first, quantity second

dr. Darko Zupanc & dr. Gašper Cankar, National Examinations Centre, Slovenia

Prezi presentation of the whole e-learning course is available at: https://prezi.com/view/TINOwDKfKjzpndcwL56K/

EDU-LAB: Project co-funded by European Union funds (ERDF, IPA).

Table of contents

I. Program for the International Assessment of Adult Competencies (PIAAC)

Discussing quality of knowledge and skills in society, the new PIAAC data allow us to directly observe the skills that people currently have, not just the formal qualifications that they once obtained. *Program for the International Assessment of Adult Competencies (PIAAC)* (http://www.oecd.org/skills/piaac/)

PIAAC is an OECD Survey of adult skills (sometimes called also 'PISA for adults') that was designed to measure adults' proficiency in several key information-processing skills, namely literacy, numeracy and problem solving in technology-rich environments.

There were two rounds of data collection: from August 2011 to November 2012 in 24 countries and economies and from April 2014 to March 2015 in another 9; all together in 33 countries and economies.

Adults who are highly proficient in the skills measured by the survey are likely to be able to make the most of the opportunities created by the technological and structural changes modern societies are going through. OECD (2016),

OECD (2016), *Skills Matter: Further Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris. *http://dx.doi.org/10.1787/9789264258051-en*

Countries greatly differ in literacy proficiency among 20-24 year-olds. In Figure 1, we show two sets of data: Mean PIAAC literacy proficiency of those, without tertiary and not enrolled in tertiary education, AND with tertiary education or enrolled in tertiary.

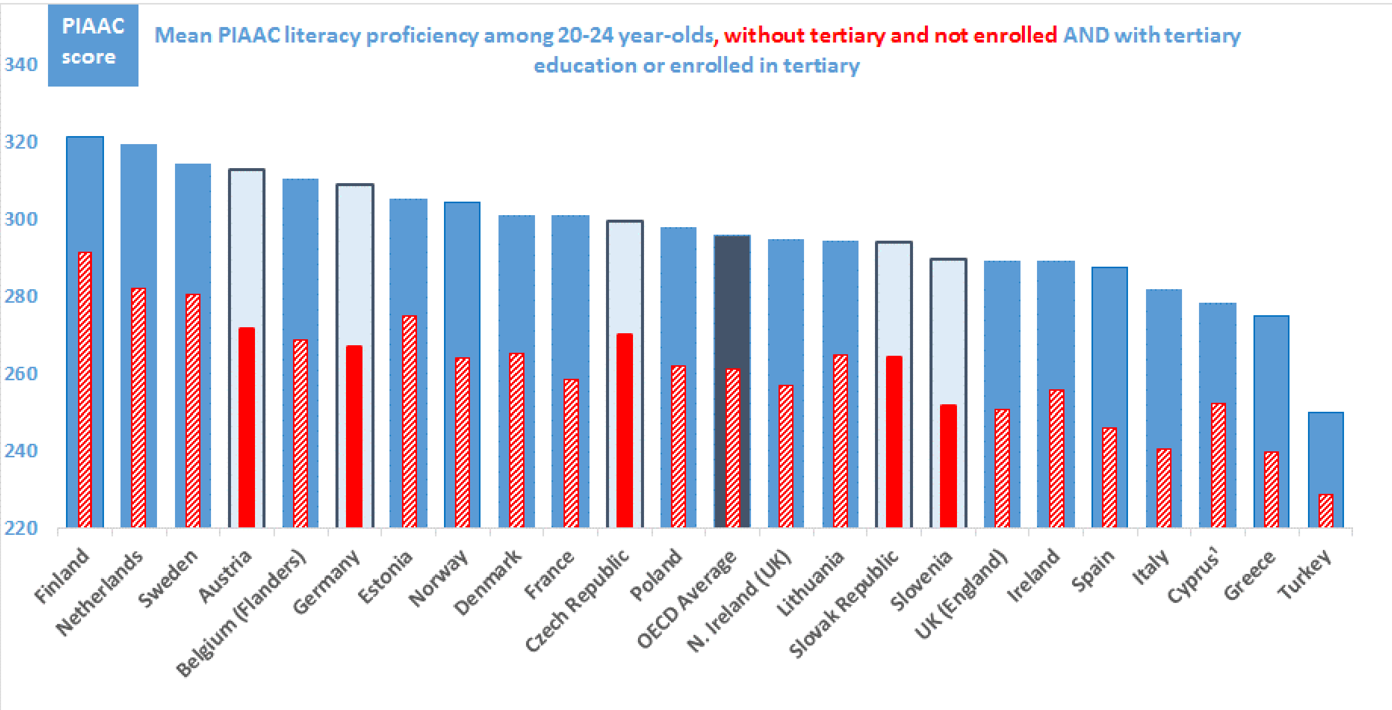


Figure 1

Adults with or enrolled in tertiary education typically demonstrate higher literacy proficiency, although in some countries the achievements of adults without tertiary are nearly identical to achievements of adults with tertiary education in other countries.

Examples:

PIAAC data shows that 25-65 year-old Italian adults with tertiary education achieve similar literacy proficiency (281.3) than Japanese adults without tertiary education (281.1).

Japanese adults with upper secondary education are more proficient in literacy (286.7) than Slovenian adults with tertiary education (285.7).

Dutch adults with upper secondary education are more proficient in literacy (283.1) than Spanish adults with tertiary education (282.0).

Dutch adults with upper secondary education are equally proficient in numeracy (280.7) as Italian adults with tertiary education (280.2).

Danish adults without upper secondary education are equally proficient in numeracy (241.0) as USA adults with upper secondary education (241.0).

Norwegian adults without upper secondary education are equally proficient in literacy (251.4) as Slovenian adults with upper secondary education (251.6).

What about the proportion of adults in tertiary education in different countries? Is there any connection between proficiency results and proportion of adults with tertiary education?

Figure 2 shows both pieces of information combined: first information is the difference of mean PIAAC literacy proficiency scores between those with tertiary or enrolled in tertiary and those without tertiary education (among 20-24 year-olds). Second information is a proportion of students in tertiary education.

Figure 2 shows no easily recognizable pattern of association between proficiency and proportion of students with tertiary education. In other words, we cannot link high proficiency results (quality) with either big or small proportion of population with tertiary education (quantity).

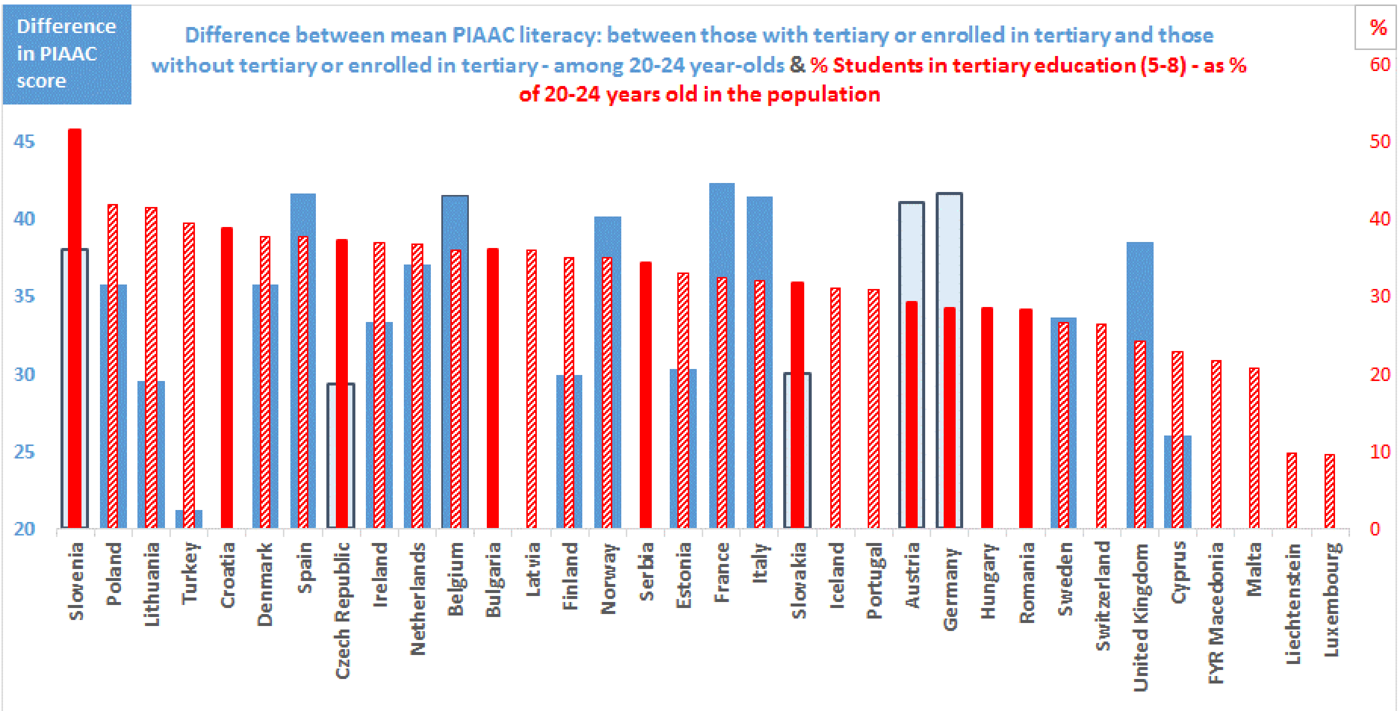


Figure 2

II. Long-lasting effects

Gustafsson (2015) suggested that PIAAC ("Programme for the International Assessment of Adult Competencies") data for representative samples of persons 16 to 65 years old combined with other data could be a source of information on differences in the quality of schooling received by different age cohorts within 33 countries and across countries. With this data, we can speculate whether the effects of education stay with a person through their lifetime into adulthood.

It is possible to control for age-effects and study the influence of other cohort effects within countries.

Gustafsson (2015) tested the idea through analyses of the Swedish PIAAC and PISA data. Performance differences between age-groups were compared with results obtained for corresponding samples in the five PISA studies and in the IEA studies (TIMSS, PIRLS…) conducted after 1990.

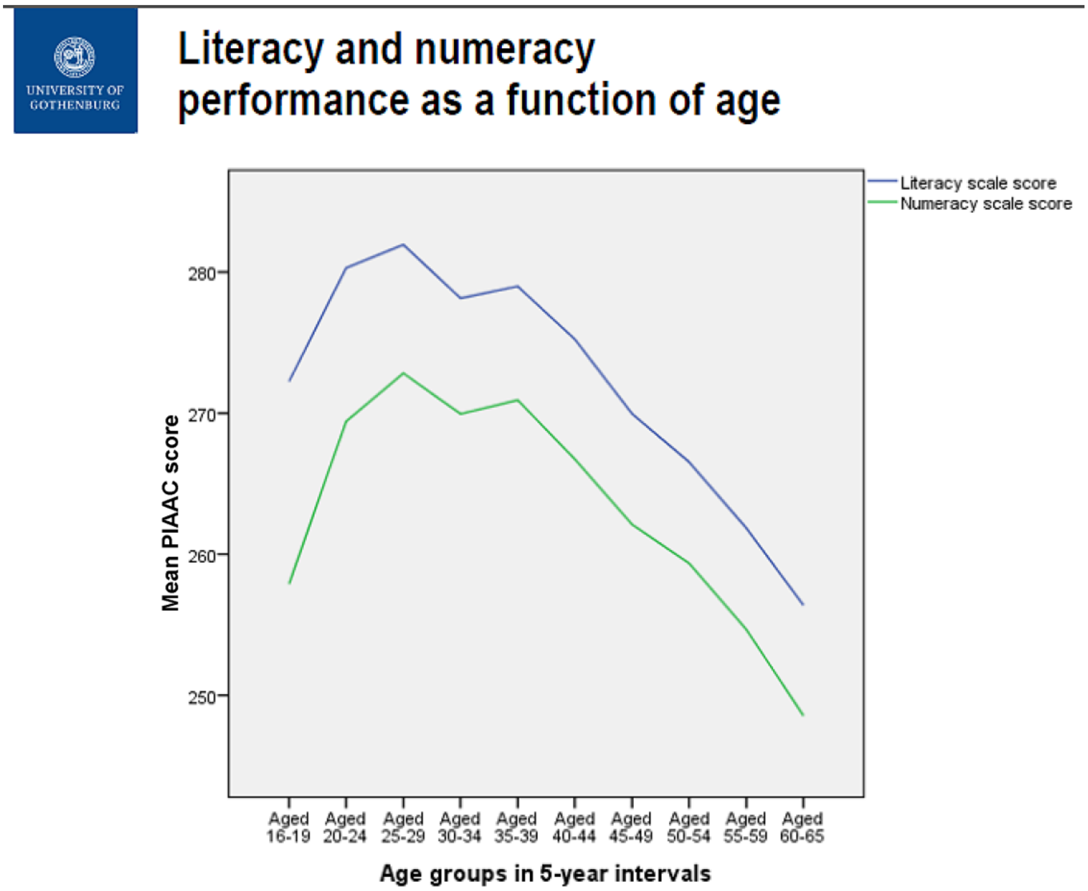
The results show good agreement between the age cohort differences in PIAAC data and the changes in level of performance in the international studies, supporting the validity of the analytical approach.

Gustafsson, J.E. (2015). *Performance Differences Between Age Cohorts as Reflectors of Differences in Quality of Education.* Key note speech in European Conference of Educational Research Network 9. Budapest, September 10, 2015.

Age-effects (Gustafsson, 2015)**:**

Performance in problem-solving tasks with new content increases up to about 25 years of age and then starts to decline. This is called 'age effect'.

Performance in tasks involving acquired knowledge and skills increases throughout life, except for old age.



Gustafsson, J.E. (2015). *Performance Differences Between Age Cohorts as Reflectors of Differences in Quality of Education.* Key note speech in European Conference of Educational Research Network 9. Budapest, September 10, 2015.

Age effects arise from individually determined factors such as maturation and decay, while **cohort effects** arise from circumstances and events that are unique to a specific age cohort, such as quantity and quality of education, nutrition and health care, and demographic factors, just to mention a few…

Cohort effect are: mass media experience, demographic factors, including migration, nutrition, health care and **quantity and quality of education** Gustafsson (2015)

Are changes in PISA results reflected in cohort differences? (Gustafsson, 2016b)

In other words – if we can see that a generation at the age of 15 shows higher literacy (compared to another generation), can we still see this difference as time goes on – when this generation is older? Gustaffson (2016b) compared PISA and PIAAC data.

Of the countries participating in PIAAC, 20 also participated in PISA 2000 – 2012 (first round).

The 15-year old students participating in PISA 2000 were 27 years old when PIAAC was conducted; the students participating in PISA 2003 were 24 years old when PIAAC was conducted, etc.

For some countries, PISA results have improved since 2000; for other countries, results have declined; and for others still, results have been stable or irregular.

Such changes might be possible to track as performance differences between PIAAC age cohorts

Gustafsson, J.-E. (2016a). Lasting effects of quality of schooling: Evidence from PISA and PIAAC. *Intelligence, 57*, 66–72

Gustafsson, J.-E. (2016b). Lasting Effects of Quality of Schooling: New Evidence from Combining PISA and PIAAC. Nordic Seminar on Basic Skills, Copenhagen November 17. University of Gothenburg

Sweden's declining average scores in PISA in each respective subject area, 2000 - 2012 and declining average scores from 1995 to 2001 for pupils in grade 8 in TIMSS mathematics.

Figure I: Sweden's average scores in PISA in each respective subject area, 2000–2012.

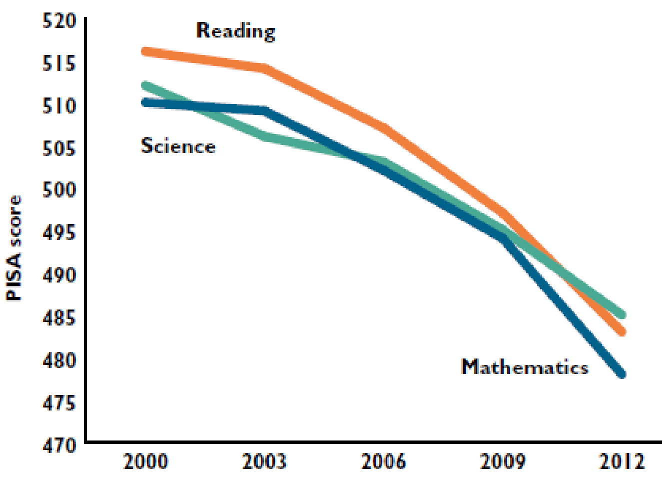


Figure II: Sweden's average scores from 1995 to 2011 for pupils in grade 8 in TIMSS mathematics.

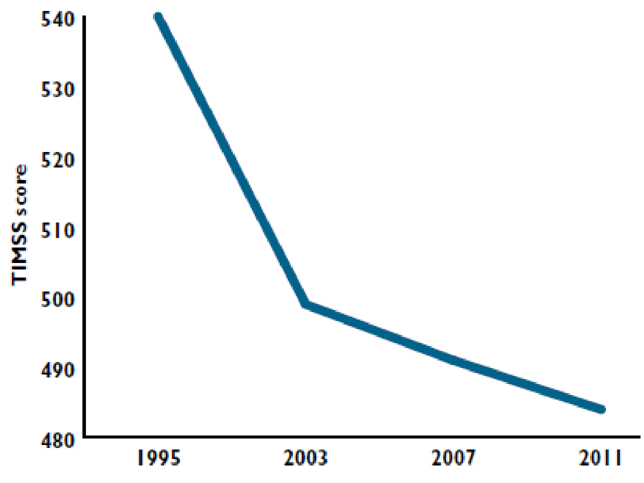


Figure I: Sweden's average scores in PISA in each respective subject area, 2000–2012. p. 10
Note: The scales for mathematics and science were only fixed in 2003 (mathematics) and 2006 (science). The scores for mathematics and science should therefore be compared with these years as reference points (Swedish National Agency for Education 2013a).
Source: Swedish National Agency for Education (2001, 2004a, 2007, 2010, 2013) and the OECD (2014).
Figure II: Sweden's average scores from 1995 to 2011 for pupils in grade 8 in TIMSS mathematics. P.18
Source: Swedish National Agency for Education (2004b, 2008, 2012a).

Henrekson, M. & Jävervall, S. (2016).

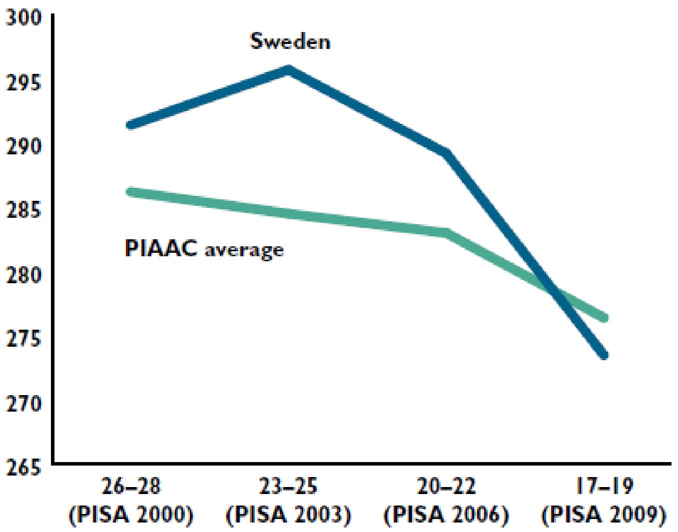
Henrekson, M. & Jävervall, S. (2016). Educational performance in Swedish schools is plummeting – what are the facts? https://www.iva.se/globalassets/info-trycksaker/iva/201611-iva-henrekson-javervall-english-f.pdf

As PIAAC assesses numeracy and literacy, these two areas are compared with those in PISA. Figures 25 and 26 show comparisons between these age groups for Sweden and the average of countries that participated in PIAAC. Comparisons between PIAAC and PISA (Henrekson & Jävervall, 2016).

The groups of pupils who participated in the four PISA assessments can be identified by the following age groups in PIAAC:

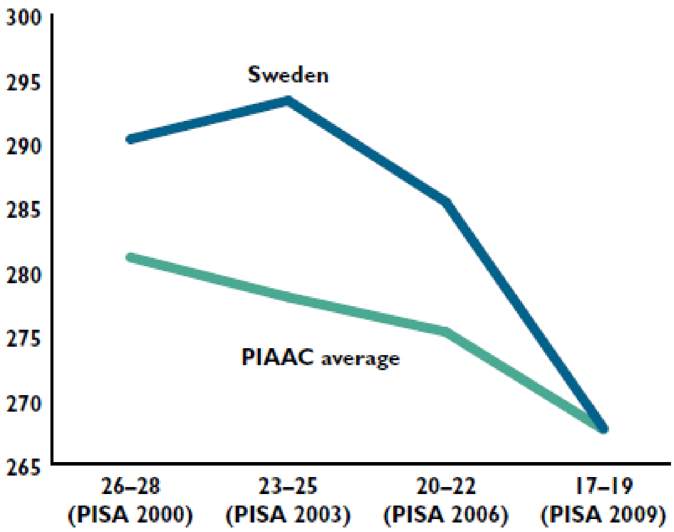
- PISA 2000: 26 - 28 years old in PIAAC (12 years later)
- PISA 2003: 23 - 25 years old in PIAAC (9 years later)
- PISA 2006: 20 - 22 years old in PIAAC (6 years later)
- PISA 2009: 17–19 years old in PIAAC (3 years later)

Figure 25: Average scores for Sweden in reading and PIAAC averages for the age groups 26–28, 23–25, 20–22 and 17–19 in PIAAC 2012.



Source: OECD (2013b).

Figure 26: Average scores for Sweden in mathematics and PIAAC averages for the age groups 26–28, 23–25, 20–22 and 17–19 in PIAAC 2012.



Source: OECD (2013b).

Henrekson, M. & Jävervall, S. (2016). P. 33

Figures 25 and 26 **mirror** Sweden's trends in the PISA assessments.

Swedish pupils, who took the PISA assessment in 2000, perform well twelve years later in PIAAC in both reading and numeracy.

Swedes who took PISA in 2003 achieved the best results overall.

For the 20 - 22 and 17 - 19 age groups, i.e. those who took the PISA tests in 2006 and 2009, average scores are lower, indicating that they were not able to compensate for deficiencies in knowledge at compulsory school through increased learning at a later stage. This applies to both literacy and numeracy.

Gustafsson (2016a, 2016b) concluded:

PISA achievement trends at age 15 are reflected in differences in the level of performance between adjacent age-groups in PIAAC

•**The level of achievement at the end of compulsory school thus lasts up** to at least age 27, perhaps even life-long (Gustafsson 2016a, 2016b & Rosdahl, 2016)

•The results also show that the differences between age cohorts carry information about quality and quantity of schooling and are affected by general societal development (changes in primary and secondary school systems)

•It is **essential that the compulsory school offers high-quality learning opportunities, because effects are lasting!**

Gustafsson, J.-E. (2016a). Lasting effects of quality of schooling: Evidence from PISA and PIAAC. *Intelligence, 57*, 66–72

Gustafsson, J.-E. (2016b). Lasting Effects of Quality of Schooling: New Evidence from Combining PISA and PIAAC. Nordic Seminar on Basic Skills, Copenhagen November 17. University of Gothenburg

Rosdahl, A. (2016, November 17th). Reading skills, learning and skills loss after compulsory education. Nordic Seminar on Basic Skills. The Danish National Centre for Social Research (SFI), Denmark.

Moreover, the Sulkunen and Malin (2017) study for Nordic countries shows that the **significance of the length and scope of the initial education in developing literacy proficiency overall is difficult to compensate**.

The role of basic education cannot be ignored even after completing secondary and tertiary degrees!

What is happening within the country in pre-tertiary education is crucial for the quality in higher education as well.

Sulkunen, S. & Malin, A. (2017). Literacy, age and recentness of education among Nordic adults. Scandinavian Journal of Educational Research. Published online: 05 Jun 2017. http://dx.doi.org/10.1080/00313831.2017.1324898

Sulkunen, S. & Malin, A. (2016, November 17th). LITERACY, AGE AND RECENTNESS OF EDUCATION. Nordic Seminar on Basic Skills 17 November 2016. University of Jyväskylä, Finland.

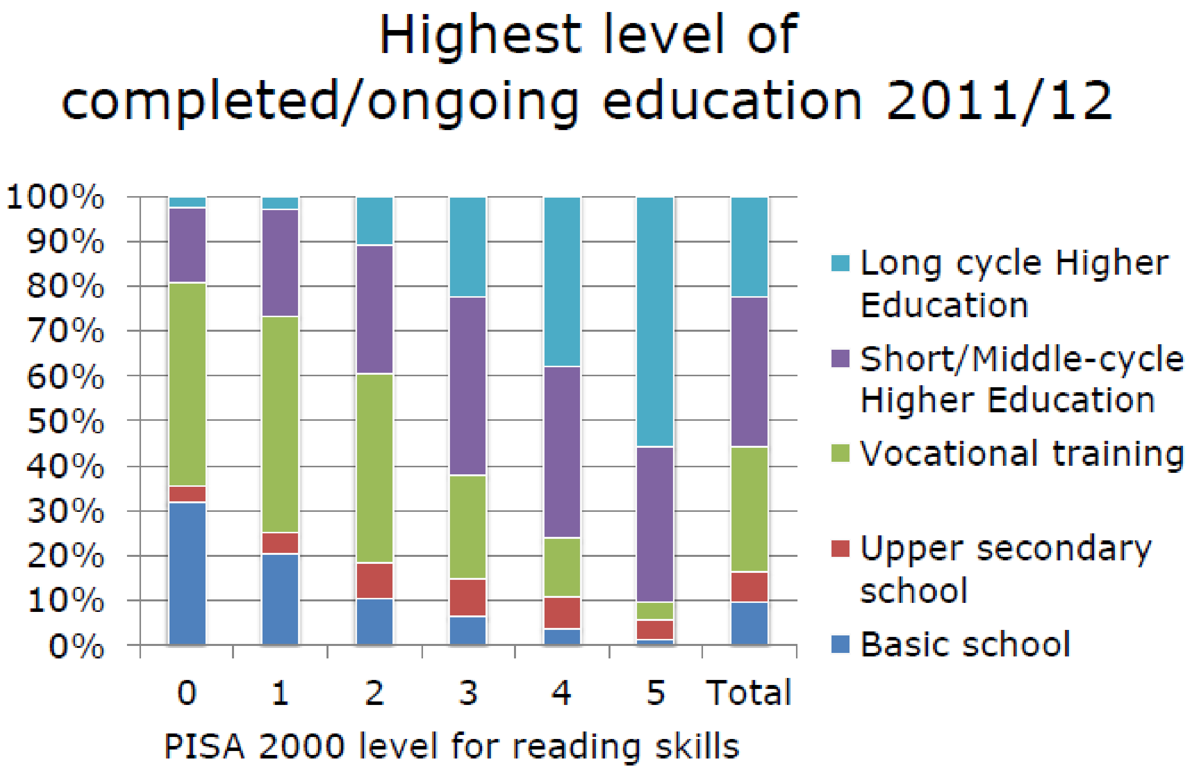
As pointed out by Hanushek (2016) with regard to adults and in the long run for society, it is really important to work on basic skills and early schooling!

Reading skills among young people are developed primarily through education. Rosdahl (2016)

In Denmark as well, reading skills at age 15 are associated with the life course of young people up until age 27.

Rosdahl (2016) shows the importance of the skill level at age 15 for the highest level of completed education. Good reading skills at age 15 lead to higher proportion of students in Higher education

Good reading skills at age 15 also have positive impact on reduced unemployment / less sickness (Rosdahl, 2016)



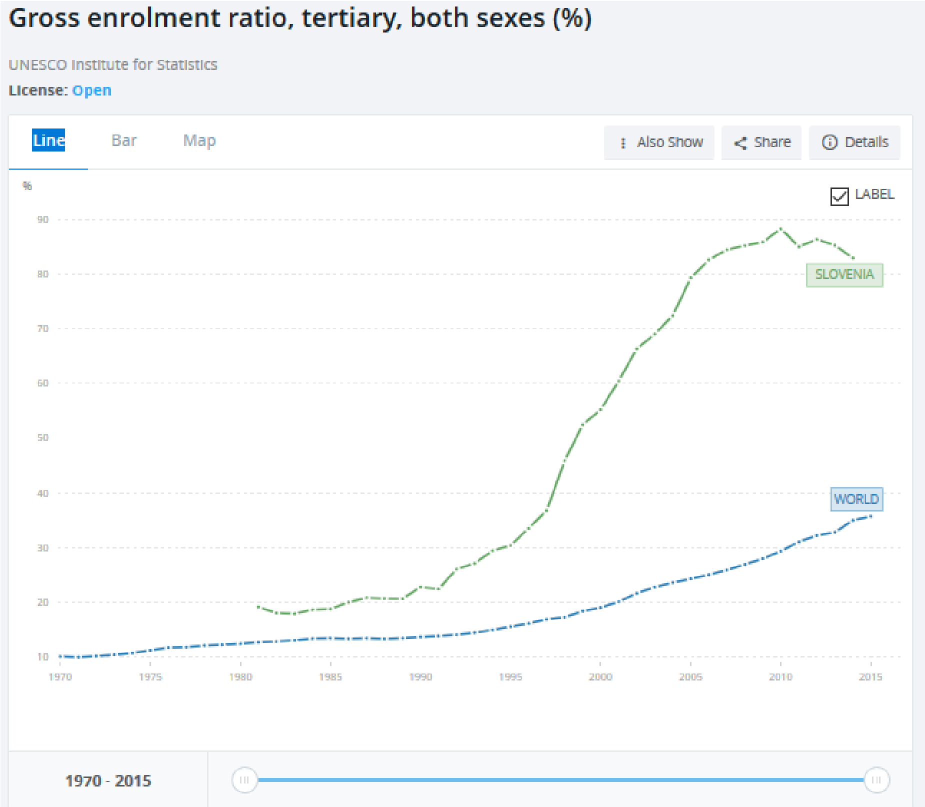
Rosdahl, A. (2016, November 17th). Reading skills, learning and skill loss after compulsory education – winners and losers in a generation of young people based on matched PISA-PIAAC data. Nordic Seminar on Basic Skills. The Danish National Centre for Social Research (SFI), Denmark.
https://dk.sfi.dk/nordic-seminar-on-basic-skills/

III. Quantity & Tertiary Education

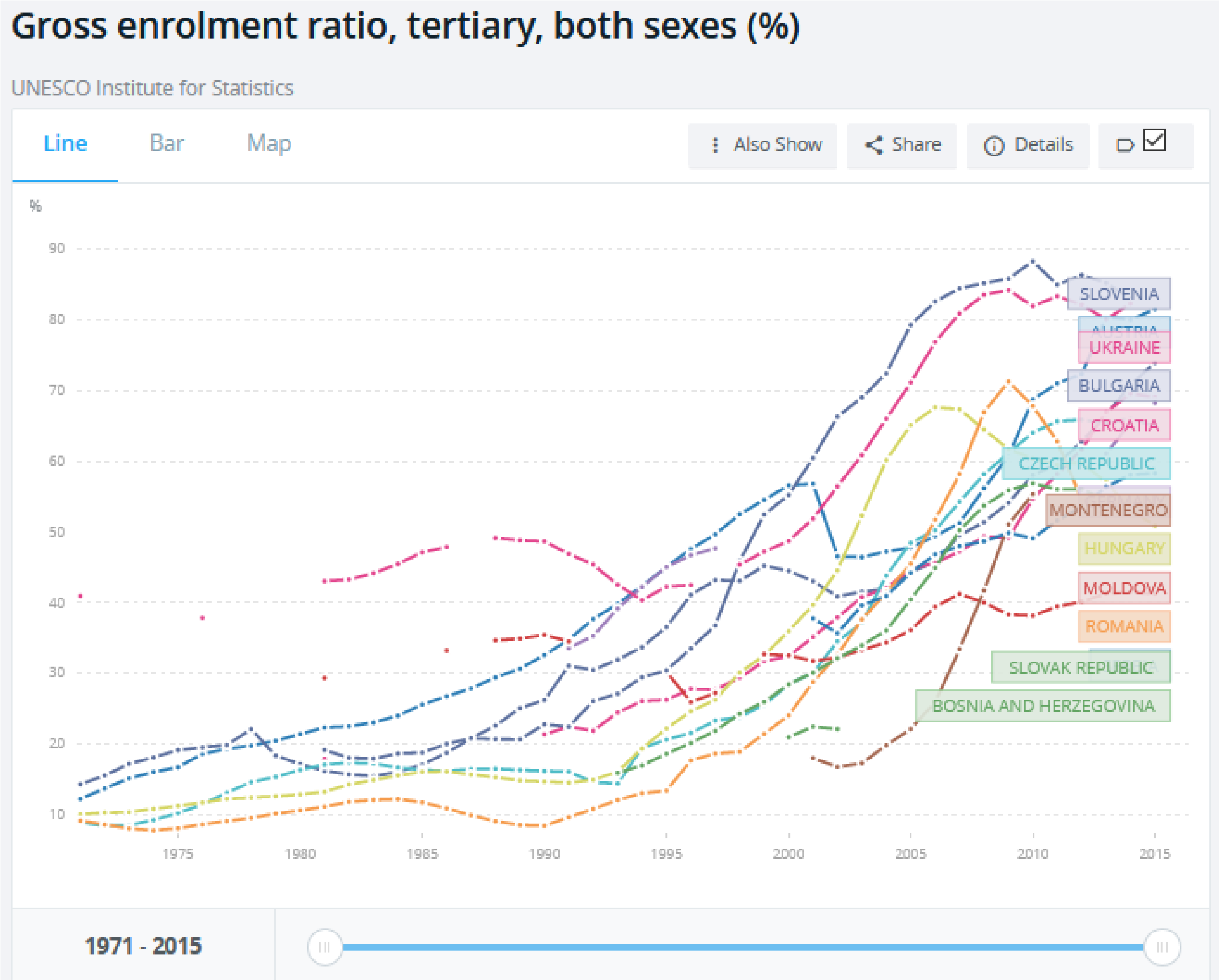
What about Quantity?

Trends in past decades...

Gross enrolment ratio, tertiary, both sexes (%); total enrolment in tertiary education (ISCED 5 to 8), regardless of age, expressed as a percentage of the total population of the five-year age group following on from secondary school leaving. *Source*: UNESCO Institute for Statistics



http://data.worldbank.org/indicator/SE.TER.ENRR?view=chart

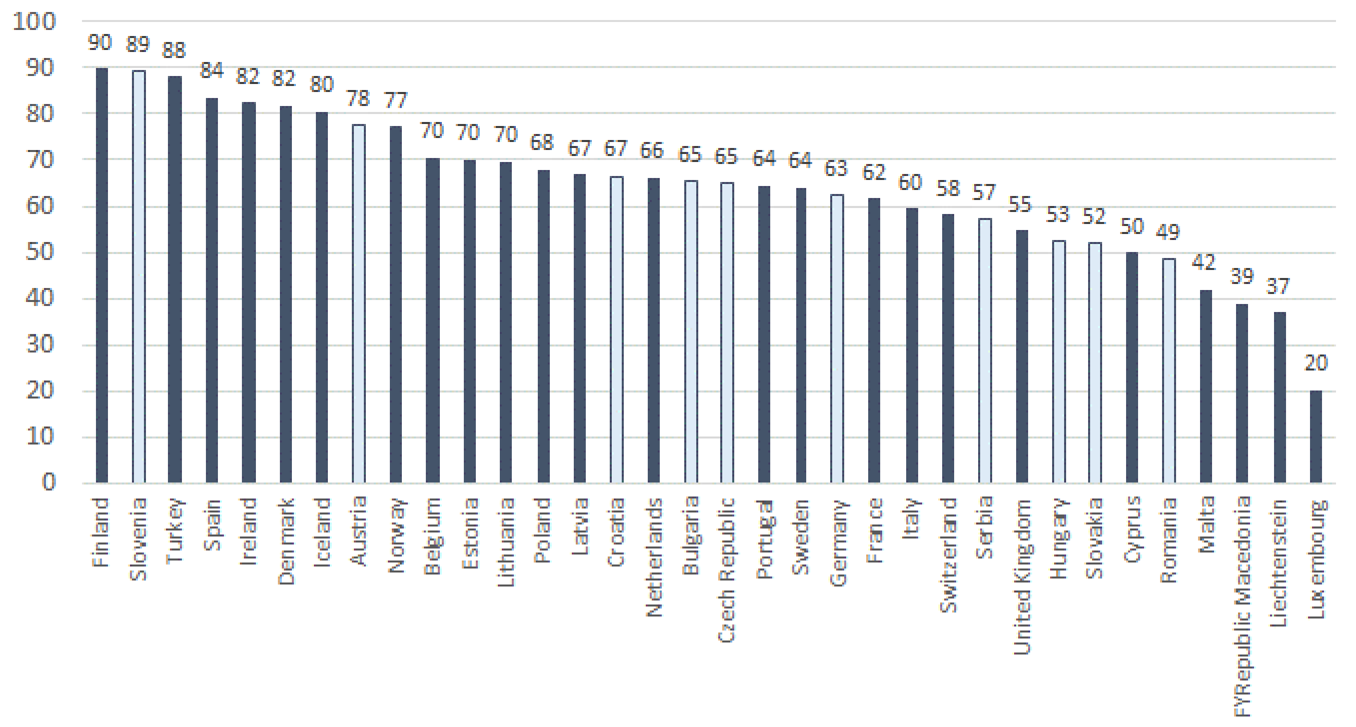


http://data.worldbank.org/indicator/SE.TER.ENRR?locations=AT-BA-BG-HR-CZ-DE-HU-MD-ME-RO-RS-SK-SI-UA&view=chart

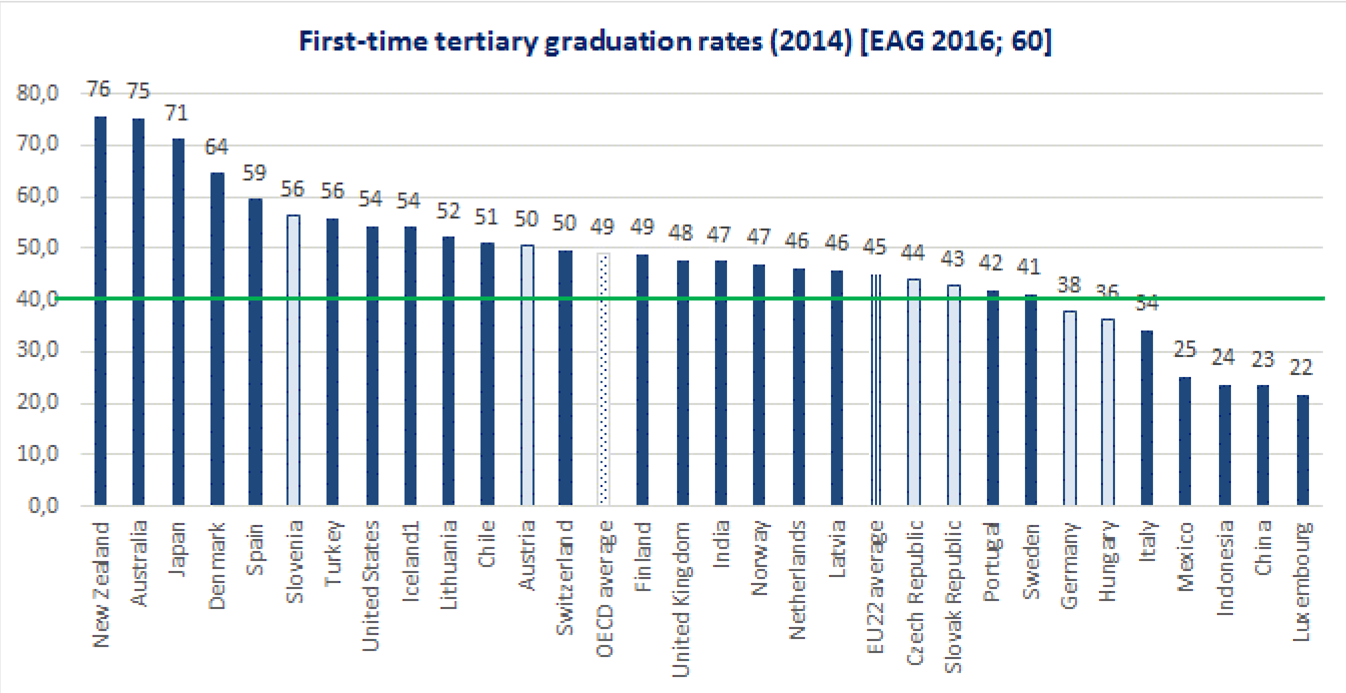
Selected Countries and Economies

Country	1971	2015
Austria	12	82
Bosnia and Herzegovina		
Bulgaria	14	74
Croatia		69
Czech Republic	9	65
Germany		68
Hungary	10	51
Moldova		41
Montenegro		
Romania	9	53
Serbia		58
Slovak Republic		
Slovenia		
Ukraine	41	

Students in tertiary education (ISCED11 5-8) - as % (age: total) in the population 2014 [Eurostat]



In the majority of countries first-time graduation rates are above 40% (The EU 2020 target for 30 to 34 years old); in some of them rates are 60 % or 70 %; the trend is growing



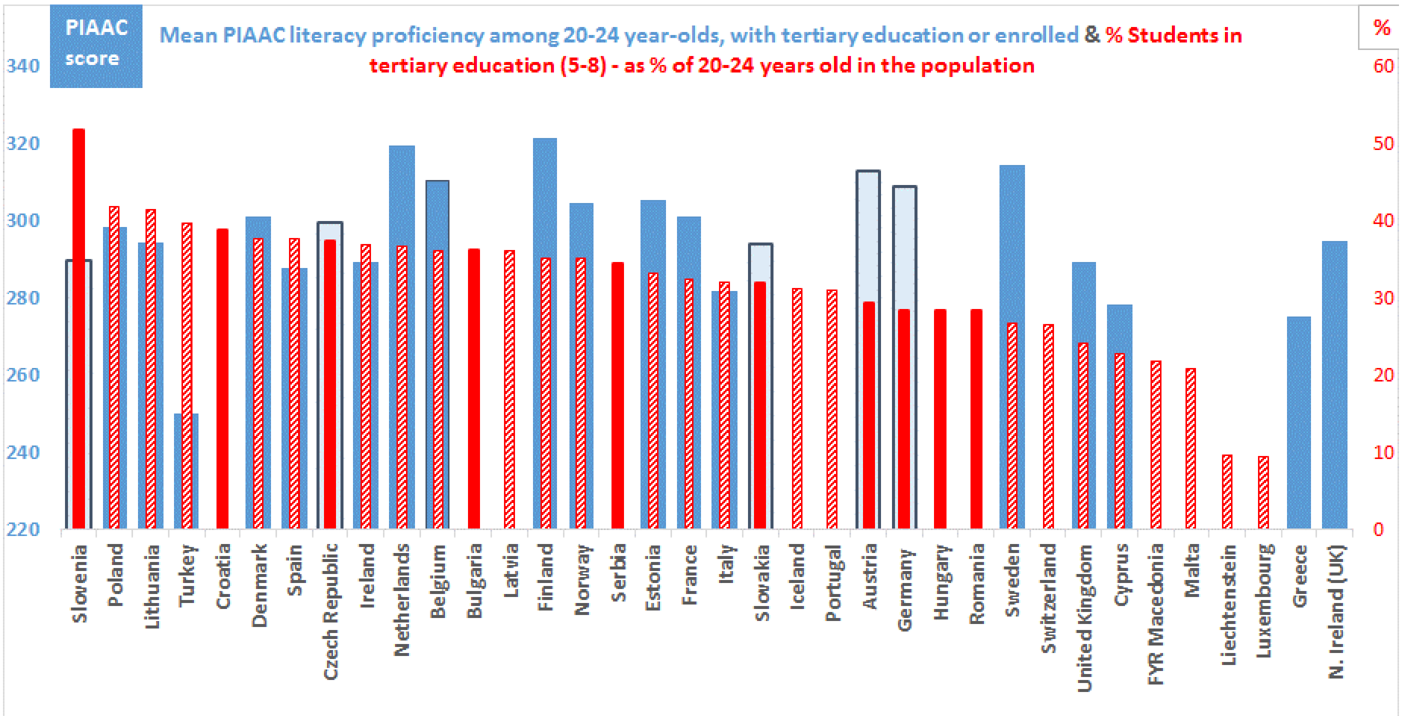
What is the mean level of PIAAC literacy proficiency in countries and economies with a different proportion of adults with tertiary attainment?

There are big differences between countries. Some countries have a low proportion of adults with tertiary attainment and high mean level of PIAAC literacy proficiency. On the contrary, some countries have a high proportion of adults with tertiary attainment and a low mean level of PIAAC literacy proficiency.

Consider the combination of two distributions for countries and economies in Europe: tertiary educational attainment distribution & PIAAC literacy proficiency for those with tertiary education or enrolled in tertiary education

For planning and governance of higher education policies, in each country and economy, generations between 20 - 24 years of age, with tertiary education or enrolled in tertiary education, are of special interest.

We cannot recognize any pattern between two sets of data.



What are the employment trends by level of attainment? EU-28, 2008, 2015, 2020 and 2025; Cedefop 2015 skills forecast.

Do the employment trends by level of attainment correspond with the proportion of students in tertiary education or first-time graduation rates?

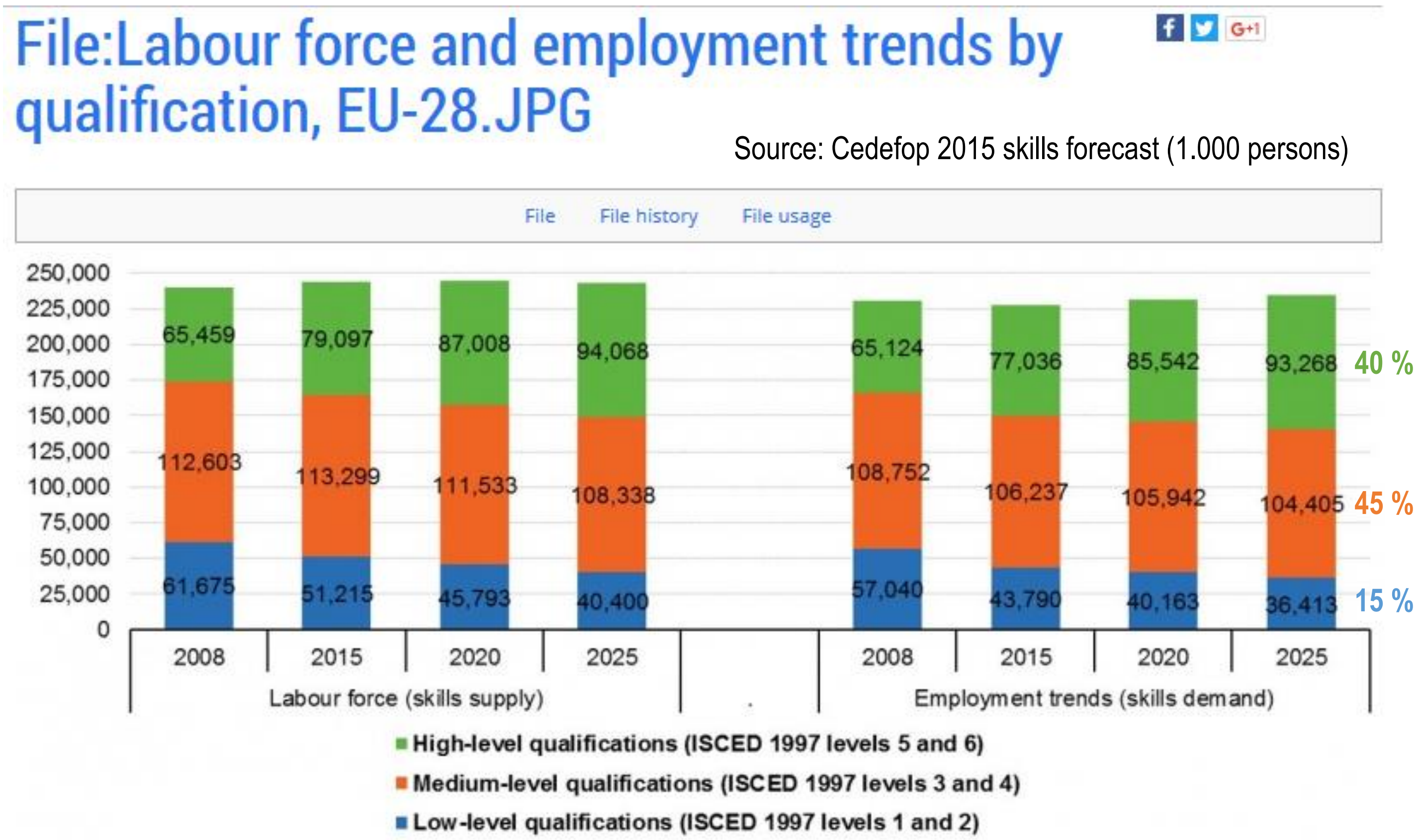


Figure: Labour force and employment trends by qualification, EU-28, 2008, 2015, 2020 and 2025 **(1.000 persons)**
http://ec.europa.eu/eurostat/statistics-explained/index.php/Europe_2020_indicators_-_employment

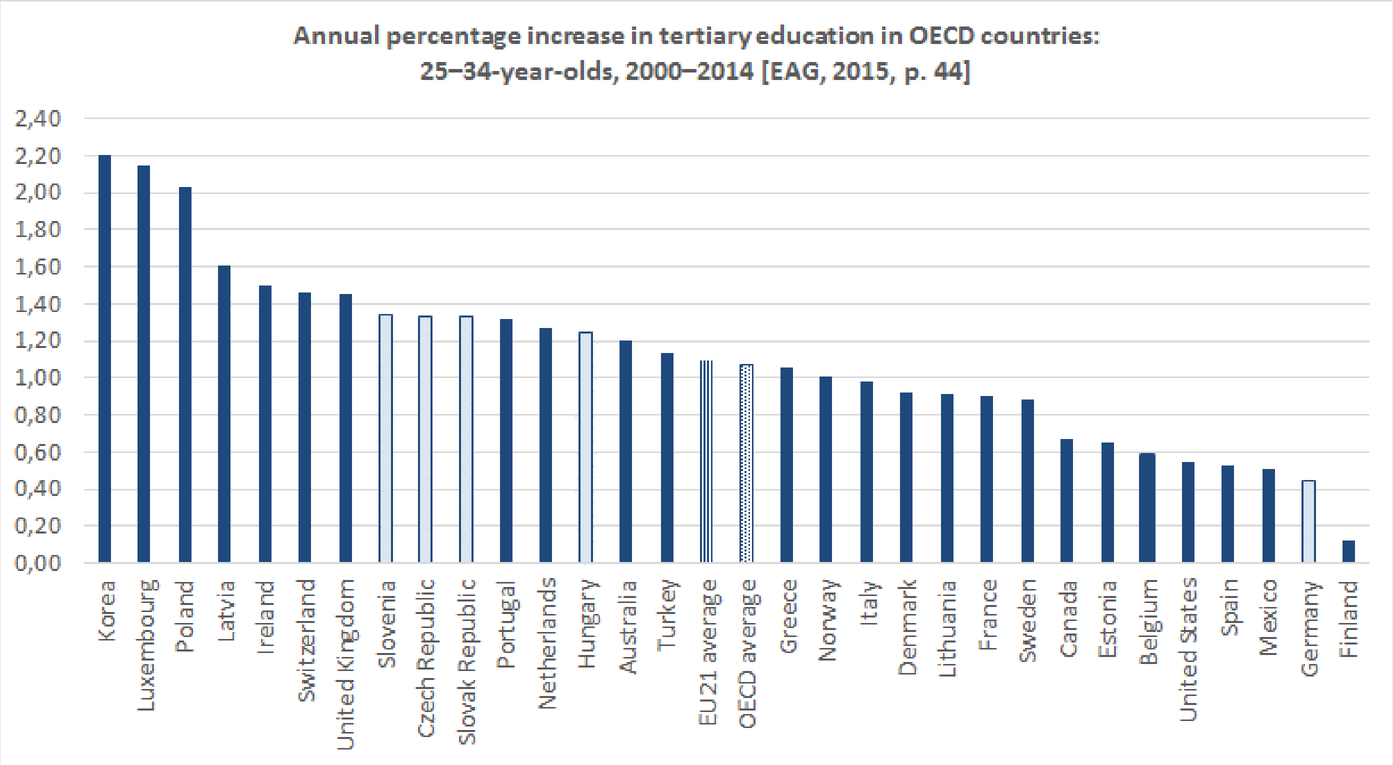
Around the world, countries have been pushing to expand education – especially at the tertiary level (Hanushek, 2016).

Higher education is seen as the source of innovation that will drive productivity improvements and thus economic growth. Moreover, the expansion of higher education is frequently put forth as an attractive government policy because of its potential impact on economic growth (e.g. Browne Report, 2010).

A variety of countries have contemplated expanding their systems of higher education.

Browne Report (2010), Securing a Sustainable Future for Higher Education: *An Independent Review of Higher Education Funding & Student Finance, 12 October*.

Eric A. Hanushek, E. A. (2016) Will more higher education improve economic growth? *Oxford Review of Economic Policy, Volume 32, Number 4*, pp. 538–552



The figures above show the annual increase in the percentage of population aged 25 - 34 with tertiary schooling over 2000 - 2014 across OECD and EU (21) countries.

The OECD average and EU (21) has increased by more than 1 percentage point per year for the last decade and a half.

This substantial increase reflects a common view that expanding higher education is also a way to promote better economic outcomes (Hanushek, 2016). When cognitive skills are accounted for, **tertiary attainment is not significantly associated with long-run growth differences across OECD countries** (Hanushek & Wössmann, 2011).

Hanushek (2016) identified the strong positive relationship between achievement levels measured by PISA test scores and overall tertiary schooling in the past, i.e. before 2000. When considering the expansion of higher education nowadays, there is no relationship with PISA 2000 scores.

The skills that can be expected at the end of university in different countries will almost certainly differ strongly; the education in college is likely to follow the earlier differences in skills. **The skills of college graduates (value added) are endogenous and depend directly on skills at entry to college**.

The policy discussion often appears to assume that the skills of college graduates are exogenous and **fixed**, implying that the expansion of higher education will lead to proportionate increases in knowledge capital.

Hanushek (2016) **suggests some potential for disappointment,** since prior growth analyses indicated that simply expanding tertiary schooling at the current quality levels is unlikely to spur new long-run growth.

Once knowledge capital, as measured by international mathematics and science tests, is taken into account, school attainment (or years of schooling) *per se* is unrelated to economic growth (Hanushek, 2016, 550).

A number of countries are following a misplaced investment strategy, if their goal is to improve economic growth. Countries might be better off investing in the margin to improve basic skills in earlier schooling (where they can be subsequently built upon in university) than simply expanding colleges and universities with existing basic skills (p. 549).

With expanding tertiary schooling, the skills of admitted students are clearly below the average skills before expansion, and expansion of students attending tertiary education is likely to lower the skills of the average graduate.

However, the achievement levels of students at an earlier age appear to provide an index of the aggregate skills of the students at the end of their schooling when each level of schooling builds on earlier knowledge.

In other words, the output of higher education should be thought of as endogenous rather than the more common view that it is exogenous to admissions. (P. 550) **The output of higher education – the skills of college graduates – depend directly on admissions standards for colleges and universities. When admission standards are high and demanding, skills of those applying for admission will increase!**

IV. Economic growth (GDP) is highly related to the knowledge capital of the country

The development of international assessments of mathematics and science TIMSS, reading (PIRLS), and different literacies (PISA), enabled an alternative approach to measure human capital directly through student achievement (Hanushek, 2016).

Economic growth is highly related to the knowledge capital of the country. Moreover, once knowledge capital, as measured by international mathematics and science tests, is taken into account, **school attainment (or years of schooling) per se is unrelated to economic growth**. With regard to the forgoing, **adding years of university provides no greater impact than added years of earlier schooling (**Hanushek, 2016)**.**

To be sure, one cannot produce electrical engineers and computer scientists without investing in higher education. However, **one can produce better engineers, if universities start with students with stronger skills**. Considering various countries, the better engineers produced in countries with greater knowledge capital appear to have a distinct impact on growth differences.

The achievement levels of students at an earlier age appear to provide an index of the aggregate skills of the students at the end of their schooling when each level of schooling builds on earlier knowledge.

Eric A. Hanushek, E. A. (2016) Will more higher education improve economic growth? *Oxford Review of Economic Policy, Volume 32, Number 4*, pp. 538–552

Quantity vs. quality

Analysing the sample of fifty countries with available international test-score data and growth information, Hanushek and Woessmann (2008) found out that quantity of schooling is statistically significantly related to economic growth in a specification that does not include the measure of cognitive skills.

… but **<u>the association between years of schooling and growth turns insignificant and its marginal effect is reduced to close to zero once cognitive skills are included in the model</u>** (see figure 8).

In other words, **<u>school attainment has no independent effect over and above its impact on cognitive skills.</u>** (Hanushek and Woessmann, 2008, 638-639)

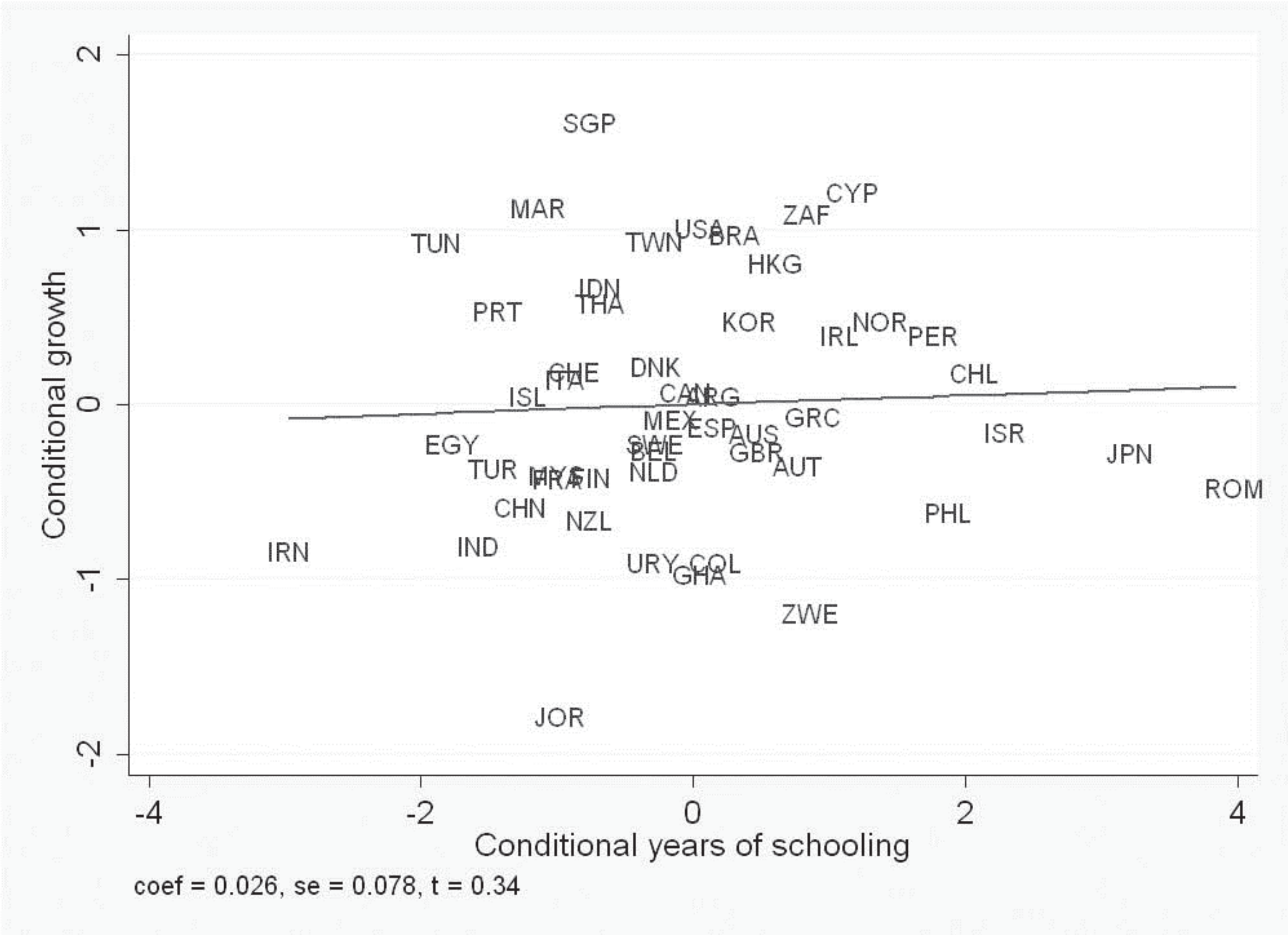


Figure 8. Added-Variable Plot of Growth and Years of Schooling with Test Score Controls

Notes: Added-variable plot of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960-2000 on the initial level of real GDP per capita in 1960, average test scores on international student achievement tests, and average years of schooling in 1960. Author calculations; see table 2, column 2.

Test scores, as opposed to years of schooling, have a powerful impact on growth.

School attainment has no independent effect over and above its impact on cognitive skills.

There is a big difference between the "quality approach" and the "access approach".

Hanushek, E. & Woessmann, L. (2008), 'The Role of Cognitive Skills in Economic Development', *Journal of Economic Literature, 46(3),* 607–68.

Returns to Quality-Adjusted Schooling: New Evidence from PIAAC

Human capital is a key driver of individual and national economic growth.

Quantity-based measures, such as years of schooling or the quality of schooling?

Schwerdt, G. (2013). **VIDEO!** from 3:15 to 8:15 minute

https://www.ets.org/c/22217/videos.html

Schwerdt, G. (2013). *Estimated Returns to Quality-Adjusted Education: New Evidence from PIAAC.* PIAAC Invitational Research Conference: The Importance of Skills and How to Assess Them. ETS & OECD, Washington, D.C., United States November 13–15.

Figure 7: After controlling for the initial level of GDP per capita and for years of schooling, **the test score measure features a statistically significant effect on the growth in real GDP per capita in 1960 - 2000**.

Test scores that are larger by one standard deviation (measured at the student level across all OECD countries in PISA) are associated with an average annual growth rate in GDP per capita that is two percentage points higher over the whole forty-year period. (Hanushek and Woessmann, 2008, 638)

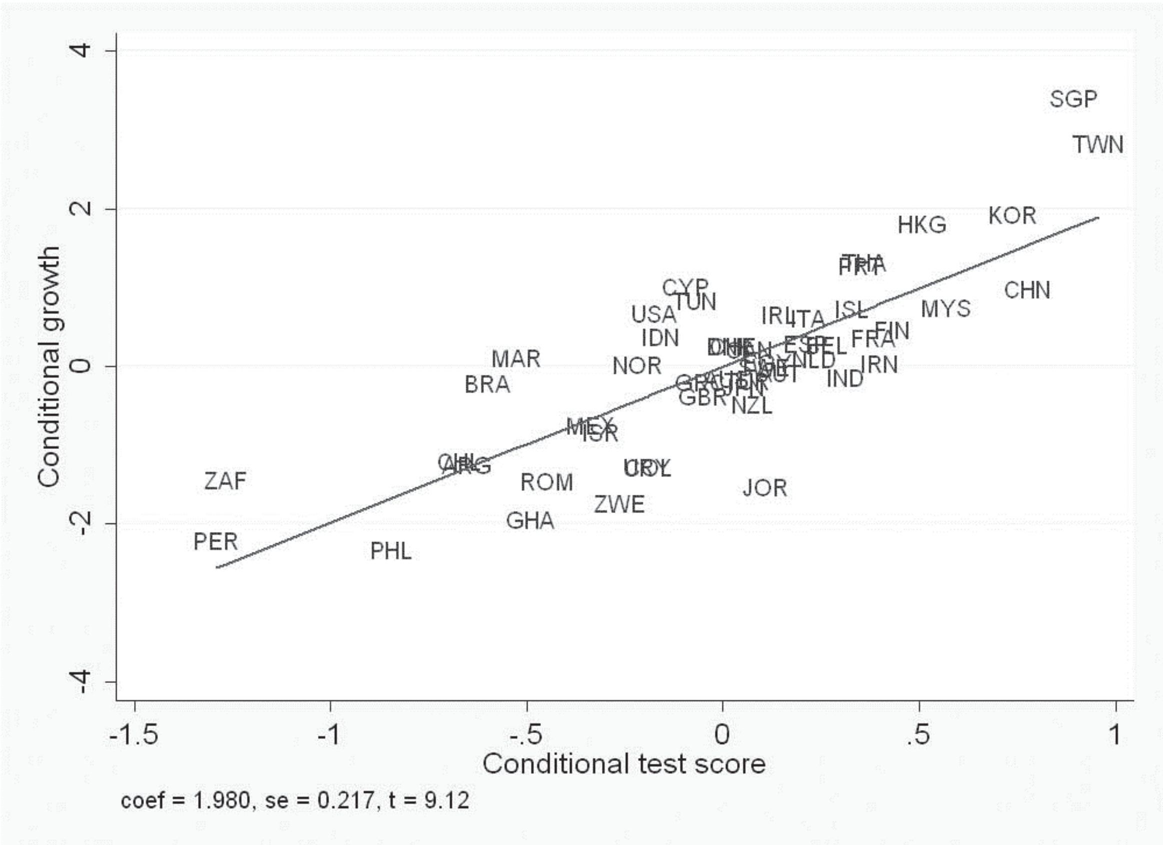


Figure 7. Added-Variable Plot of Growth and Test Scores

Notes: Added-variable plot of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960–2000 on the initial level of real GDP per capita in 1960, average test scores on international student achievement tests, and average years of schooling in 1960. Author calculations; see table 2, column 2.

The commonly used quantitative measure of years of schooling became insignificant (once cognitive skills are taken into account). Average scores on the international achievement tests between 1964 and 2003 account for more than three-quarters of the variation in long-term growth rates of per capita GDP (Woessmann, 2016).

What this means is that years of schooling in recent findings (Hanushek and Woessmann, 2015) impact GDP only as much as they are rough measures of cognitive skills of an individual. Once we account for cognitive skills, years of schooling have no effect. Lengthening the schooling (with no added quality) does not influence GDP.

Education affects go well beyond education; they affect health of children, people's own health, crime, and citizenship (e.g. Lochner (2011)).

Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives, 30(3),* 3–31.
Lochner, L. (2011). "Nonproduction benefits of education: Crime, health, and good citizenship." In *Handbook of the Economics of Education, Vol. 4*, edited by E. A. Hanushek, S. Machin, & L. Woessmann, 183-282. Amsterdam: North Holland.
Wößmann, L. (2016). "The Economic Case for Education", *Education Economics 24 (1),* 3-32,
Hanushek, Eric A., and Ludger Woessmann. 2015. *The knowledge capital of nations: Education and the economics of growth*. Cambridge, MA: MIT Press.

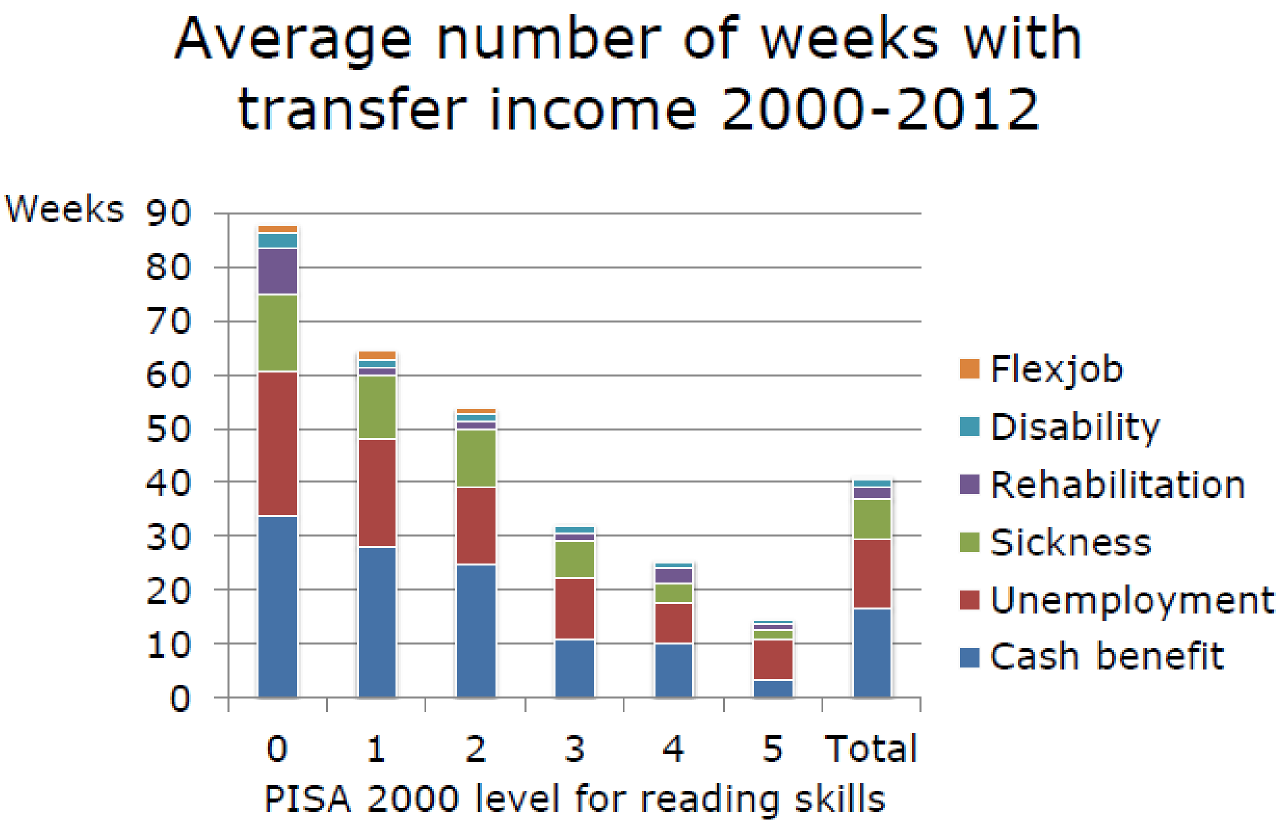
Cognitive skills are an important force in economic development
The simple conclusion from the combined evidence is that differences in cognitive skills – measured across numeracy, literacy, and problem-solving domains – lead to economically significant differences in economic growth. **School policy can, if effective, raise cognitive skills** (Hanushek & Woessmann, 2012).
One standard deviation higher cognitive skills of a country's workforce is associated with approximately two percentage points higher annual growth in per capita GDP.
Aggregate data provide direct evidence that both providing broad basic education — education for all— and pushing significant numbers to very high achievement levels have economic payoffs.
If the EU were successful in improving average student achievement by the equivalent of 25 PISA points, the economic gain would amount to an astounding €35 trillion until 2090. Put differently, this amount is the cost to the EU of not improving the quality of its school systems.
Policy-makers interested in advancing future prosperity should focus on educational outcomes, rather than just inputs or length of study.
Similar measures appear promising for the European system of higher education (Woessmann, 2014).

Hanushek, E. A. & Woessmann, L. (2012). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation . *Journal of Economic Growth,* 17(4), December 2012, pp. 267-321.
Hanushek, E. A., Schwerdt, G., Wiederhold, S. & Woessmann, L. (2015) Returns to Skills around the World: Evidence from PIAAC. *European Economic Review,* 73, pp. 103-130.

V. Achieved cognitive skills and proficiency in literacy skills impact non-educational and non-economic features as well

Educational outcomes and economic growth, risk of unemployment, individual earnings …
Knowledge and skills of the population are a leading determinant of economic growth, employment, and earnings in modern knowledge-based economies such as the EU (Woessmann, 2014).
Better education is very closely linked to individual and societal prosperity as reflected in earnings, employment, and economic growth.
Higher cognitive skills are systematically related to higher wages in all 23 participating countries in the first round of PIAAC (Hanushek, Schwerdt, Wiederhold, & Woessmann, 2015).
Education reduces risk of unemployment.
Education raises individual earnings: Among those who have a job, earnings increase on average by 7.4% for each additional year of education or, when measured directly as skills in PIAAC, by 17.4% for each step on the five-step competency scale.
Workers with higher PIAAC proficiency in information processing skills **are more likely to be employed** (in some countries), **earn higher wages** (in most countries) and **have better social outcomes** (in all countries).

Woessmann, L. (2014). *The Economic Case for Education*. EENEE Analytical Report 20, Dec. 2014, http://www.eenee.de/eeneeHome/EENEE/Analytical-Reports.html



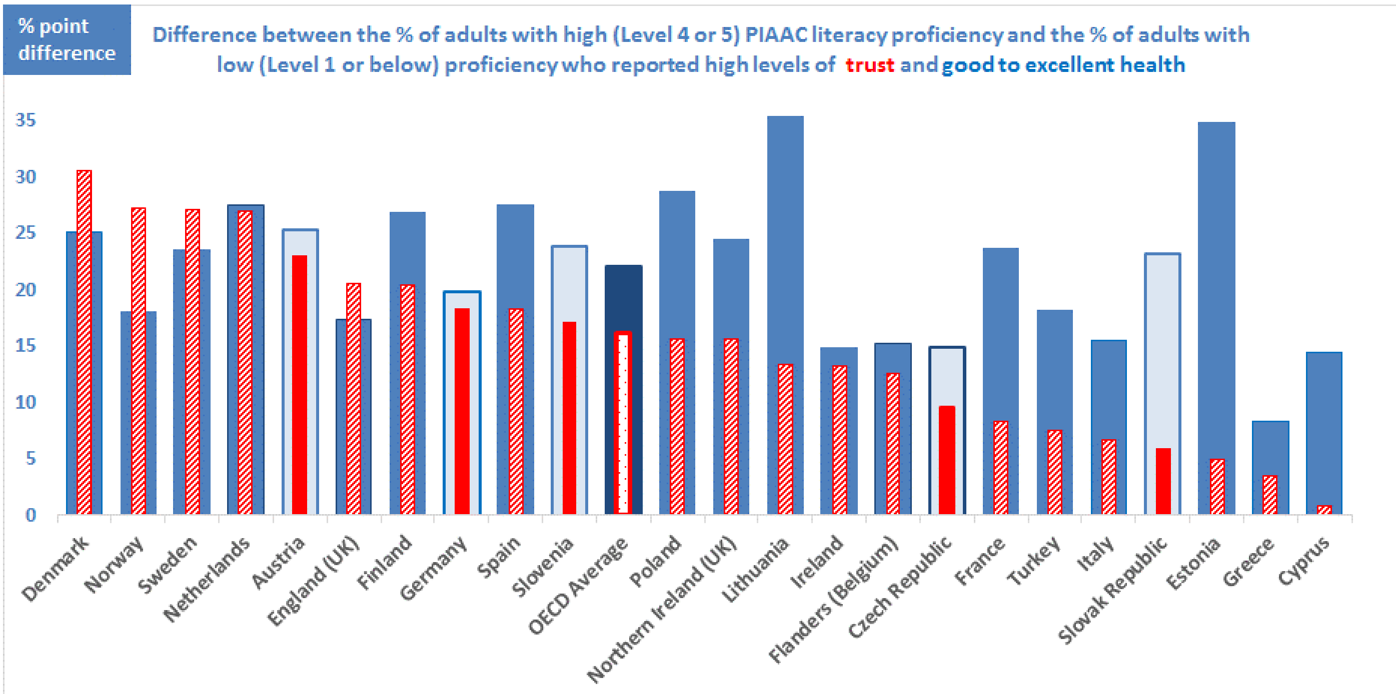
TRANSFER INCOME definition: money from the government in the form of benefits (= payments for people who cannot find a job or are too ill to work),.

Rosdahl, A. (2016, November 17th). Reading skills, learning and skill loss after compulsory education – winners and losers in a generation of young people based on matched PISA-PIAAC data. Nordic Seminar on Basic Skills. The Danish National Centre for Social Research (SFI), Denmark.
https://dk.sfi.dk/nordic-seminar-on-basic-skills/

While employability and wages are important for the individual's well-being, individuals and policy makers are becoming aware that non-economic factors also contribute to individual well-being and to the smooth functioning of societies as a whole (OECD, 2016, 146)

The PIAAC Survey of Adult Skills collects information on four non-economic outcomes: Level of trust in others; participation in associative, religious, political, or charity activities (volunteering); sense of being able to influence the political process (i.e. political efficacy); and self-assessed health conditions.
Across countries and economies, there is a positive correlation between skills proficiency in literacy and trust, volunteering and political efficacy (with correlation coefficients in the order of 0.40).

OECD (2016), *Skills Matter: Further Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris.
http://dx.doi.org/10.1787/9789264258051-en



Interpersonal trust, especially generalized trust, is a strong predictor of economic prosperity and individual well-being. Generalized trust develops out of a feeling of goodwill towards anonymous others and enables smooth social and economic interactions in complex societies where people engage frequently in interactions with others, whom they do not know and from whom they differ in many ways. In such contexts, the absence of trust can result in negative consequences for economic activity (OECD, 2016, 144).
Across countries and economies, there is a positive correlation between skills proficiency in literacy and trust.

Health is a crucial element of individual well-being, as well as an area that absorbs a significant share of public expenditure. Increasingly complex healthcare systems, requiring adults to process a large amount of health-related information, as well as increased polarization by which highly skilled individuals are more and more likely to end up in "good" jobs and to be able to afford living in "good" places, all strengthen the link between health and proficiency in information processing skills (OECD, 2016, 147).
The relationship between health and literacy proficiency is strong in most countries/economies that participated in the PIAAC Survey of Adult Skills. On average, the chances of reporting good to excellent health are 22 percentage points higher among people who scored at Level 4 or 5 than among those who scored at or below Level 1.

OECD (2016), *Skills Matter: Further Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris. *http://dx.doi.org/10.1787/9789264258051-en*

VI. Research findings and some lessons from best performers

Measure of Quality of education - to measure human capital through student achievement

Many traditional policies of simply providing more funds for schools or adding specific resources such as smaller classes do not provide much hope for significant improvements in student achievement.

Policy makers should focus on the outcomes for individuals in terms of skills not just the inputs as the level of education in terms of years of schooling (Hanushek, 2016).

A growing body of research shows that teacher quality is the primary driver of student achievement; however, differences in quality are not closely related to teacher education and experience (Hanushek, Wössmann, 2011).

A substantial part of the cross-country variation in student achievement corresponds with differences in **external exams**, **school autonomy**, private competition and **early tracking** (Woessmann, 2016).

When cognitive skills are accounted for, tertiary attainment is not significantly associated with long-run growth differences across OECD countries (Hanushek & Wössmann, 2011)

Hanushek, E. A. (2016) Will more higher education improve economic growth? *Oxford Review of Economic Policy, Volume 32, Number 4*, pp. 538–552

Woessmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives, 30(3),* 3–31.

Selected lessons from best performers (Andreas Schleicher, PISA, OECD)

Commitment to universal achievement; **universal educational standards** and personalization as the approach to heterogeneity in the student body (as opposed to a belief that students have different destinations to be met with different expectations and selection/stratification as the approach to heterogeneity).

A **well-established delivery chain** through which curricular goals translate into instructional systems, instructional practices and student learning (**intended, implemented and achieved curriculum**).

Clear articulation of **who is responsible for ensuring student success and to whom.**

Aligned system incentive for students: **students have incentives to take tough courses and study hard.**

Have clear and ambitious **goals** that are shared across the system and **aligned with high-stakes gateways** and instructional systems.

Gateways affect the strength, direction, clarity and nature of the incentives **influencing students at each stage of their education.**

Provide **a data-rich school environment** to combat inequities. "*Without data, you are just another person with an opinion.*" (William Edwards Deming)

The results from PISA show higher levels of socio-economic equity in school systems that **use achievement data to make decisions about the curriculum and track achievement data over time** (OECD 2010, 47).

VII. External exams

"PISA-shock" in Germany

Poor performance of German students in TIMSS 1995 and PISA 2000 gave rise to an intense political discussion about the need to reform the German school system, including the assessment system at the end of lower and upper secondary education.

Students in federal states in Germany with centralized exit examinations clearly outperformed students in other federal states. PISA 2003 results showed that students from federal states with centralized examinations performed better than students from any state without centralized examinations (Wößmann 2010).

In this respect, the tendency in the German education policy was that federal states have increasingly adopted centralized exit examinations. While until 2002, only some federal states had centralized exit examinations at the end of secondary education, in 2008 only a single state had no centralized exit examinations at the end of secondary education (Wößmann 2008, p 824).

Across OECD countries, those that use **Curriculum-Based External Examination Systems (CBEES)** tend to perform higher, even when accounting for national income: students in school systems that use CBEES perform, on average across OECD countries, 16 points higher than students in school systems that do not use these examinations (OECD 2010, p 46).

Standards-based external examinations are defined (Bishop 1998) as CBEES, if they offer signals of student accomplishment that have real consequences for the student and define achievement relative to an external standard, not relative to other students in the classroom or the school.

To enable fair comparisons of achievement across schools and across students at different schools, a CBEES is organized by discipline and keyed to the content of specific course sequences, which allocates the responsibility of preparing the student for particular examinations to one or a small group of teachers.

CBEES signals multiple levels of achievement in the subject rather than indicating merely a pass-fail signal, and covers almost all secondary school students (Bishop 1998).

Between academic years 1997/98 and 2010/11, comparisons of characteristics of certified assessment at the end of general upper secondary education (ISCED 3) were done for European countries (Eurydice: *Key Data on Education in Europe* – 1999/2000 and 2012).

In 14 out of 31 countries, there were shifts for final grading at the end of general upper secondary education (ISCED 3) from students' teachers-based criteria to a more external system.

The Slovene school system is historically linked to German-speaking countries. The Matura Examination at the end of upper secondary education came to Slovenia in 1995 from English-speaking regions.

The Matura Examination at the end of upper secondary education in Slovenia is a CBEES.

With establishing CBEESs, many European countries were overtaken by Slovenia.

Benefits of the external exit examination systems

Evidence from several international student achievement tests shows that students perform substantially better in countries with external exit examinations than in countries without them (Wößmann 2008).

Positive effects of external exit examinations on test-score outcomes have been shown for Canadian provinces, for US states and for German states (Piopiunik, Schwerdt, Wößmann 2012).

Research results (Jürges, Shneider, Büchel 2003) suggest that centralized examinations increase student achievement by about one third school year equivalent.

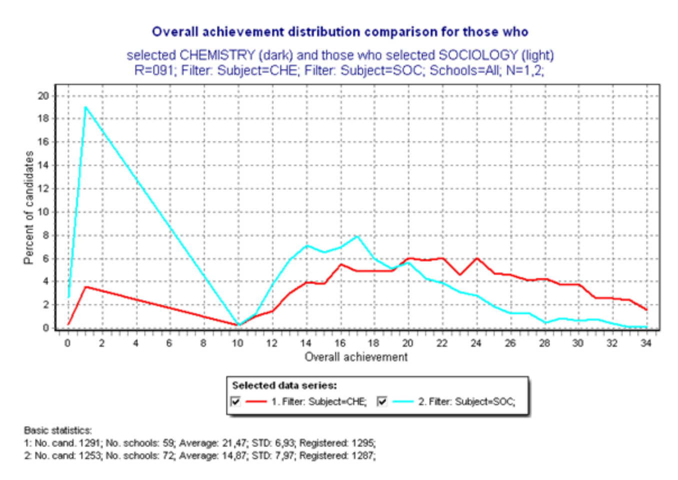
International cross-country evidence suggests that the effect of external exit examinations on student achievement may well exceed a whole grade-level equivalent (Piopiunik, Schwerdt, Wößmann 2012).

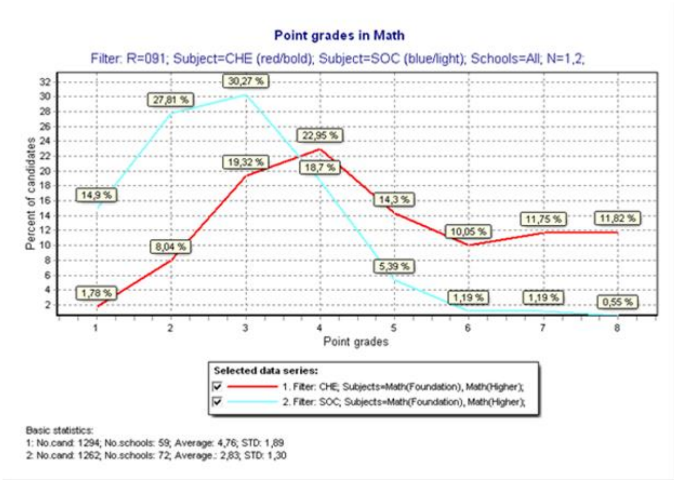
Central exit exams in Germany show 11.5 percent higher earnings and 4.3 percentage points lower unemployment probability for graduates from low-track high schools, as well as 3.2 percentage points lower unemployment probability for graduates from high-track high schools (Piopiunik, Schwerdt, Wößmann, 2013).

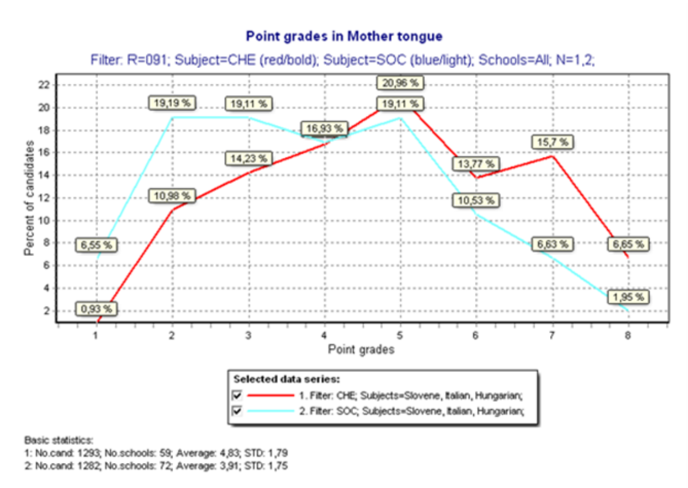
Piopiunik, M., Schwerdt, G., & Wößmann, L. (2013). Central school exit exams and labor-market outcomes. In: *European Journal of Political Economy, Vol. 31:* pp. 93-108.

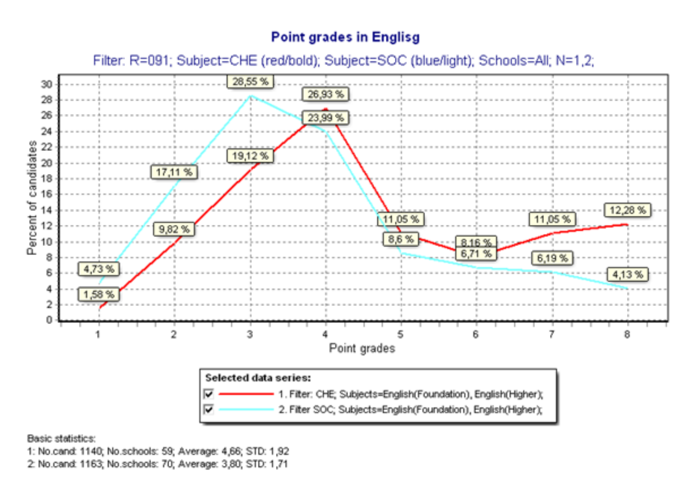
1. Higher Achievers choose certain subjects

Comparison for candidates who select different optional Matura subject: CHEMISTRY or SOCIOLOGY
Overall achievement distributions & Point grades in Math, Mother Tongue and English distributions







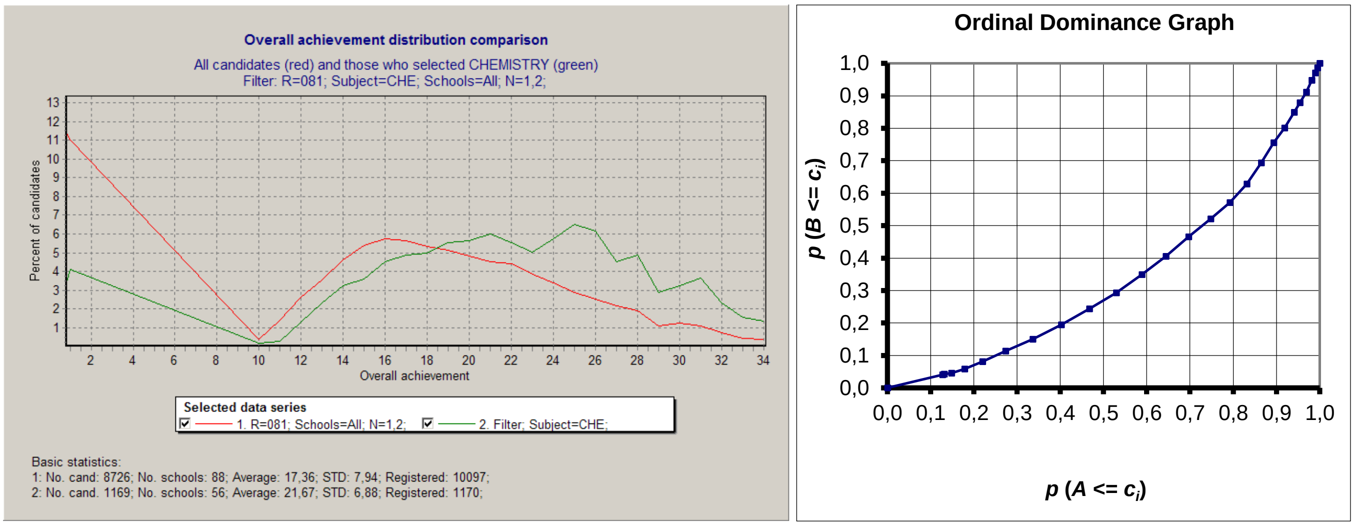


Teacher grades, Matura grades, overall achievement results and other similar educational scales have only ordinal justification - not interval. In school systems with numerical grades, means are often calculated, however, we will have to do with ordinal scales ranking and calculating medians (Nunnally & Bernstein, 1994).
One of the best-known non-parametric significance tests is the Mann-Whitney U test, also called the Mann-Whitney-Wilcoxon (MWW). The U test remains the logical choice when data are not interval but ordinal. It is often recommended for situations where the distributions of the two samples are very different. With small samples, the distribution is tabulated, but for samples above 20, there is a good approximation using the normal distribution.
In educational grade distributions, the proportion of ties is quite large; thus, we need to apply a correction for ties. (Bren & Zupanc, 2008; Siegel & Castellan, 1988). If the rankings of two groups are different, the MWW test will give us the answer. If the ranking of two groups differs, the next question is how much (Zupanc & Bren, 2009).

Zupanc, D., & Bren, M. (2009). Optional subjects in the Matura exam: Are candidates who select science subjects higher achievers? In: Conference papers: Brisbane 35th annual conference (2009), 35th Annual Conference Assessment for a creative world, Brisbane, 13.-18. September 2009. Brisbane, 2009: International Association for Educational Assessment, pp. 1-11.

http://webcache.googleusercontent.com/search?q=cache:s0RES6rvalAJ:www.iaea.info/papers.aspx%3Fid%3D77+&cd=1&hl=sl&ct=clnk&gl=si..

Overall achievement distributions for all Matura candidates (red) in year 2008 and for those who select science optional subject (green): CHEMISTRY is presented in Figure I.



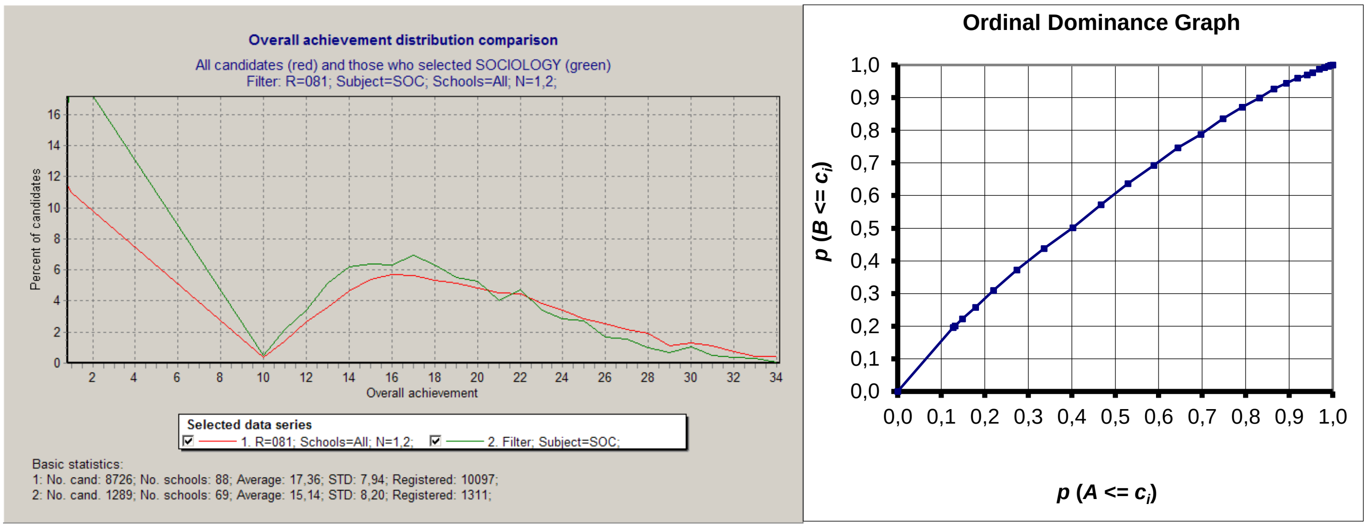
CHEMISTRY	*N*	*U*	*z*	*p*	*ρ*	*d*
A *All*	8726	3414921	-18,41	0,0000000000000	0,335	-0,330
B *CHEMISTRY*	1169	6785773	18,41	0,0000000000000	0,665	0,330

Figure I: Overall achievement distribution comparison and statistics for ALL (red) candidates and those who select CHEMISTRY (green).

http://www.iaea.info/documents/paper_4d73f71.PDF.

In the CHEMISTRY case, the MWW statistics *U is* statistically significant (p << 0,001) and the value ρ_{CHE} = 0,665 in d_{CHE} = 0,330. The index ρ area above the ODG equals the probability that a randomly chosen candidate with an optional subject (CHE) will have higher overall Matura achievement than a randomly chosen Matura candidate (A_{All}), plus one half the probability that they will have the same overall achievement. The probability for CHEMISTRY is ρ_{CHE} = 0,665. The curve is convex and stays entirely below the diagonal. The proportion of group *B* exceeds that of group A_{All} at any point on the scale. Candidates who select natural science subject (CHE) – the group *B* completely dominates all Matura candidates – group A_{All}.(Zupanc & Bren, 2009).

Overall achievement distributions for all Matura candidates (red) in year 2008 and for those who select social science optional subject (green): SOCIOLOGY is presented in Figure II.



SOCIOLOGY	***N***	***U***	***z***	***p***	***ρ***	***d***
***A** All*	8726	6498476	9,05	0,0000000000000	0,578	0,156
***B** SOCIOLOGY*	1289	4749338	-9,05	0,0000000000000	0,422	-0,156

Figure II: Overall achievement distribution comparison and statistics for ALL (red) candidates and those who select SOCIOLOGY (green).

The graph shows that candidates who select social sciences (SOC) have lower overall achievements. The MWW statistics *U* is statistically significant (*p* << 0,001). Ordinal dissimilarity between two distributions – *d* is negative and the index ρ - area above the diagonal - is less than 0.5: ρ_{SOC} = 0,422 and d_{SOC} = - 0,156. The probability for SOCIOLOGY is ρ_{SOC} = 0,422. The curve is concave and stays entirely above the diagonal. The proportion of group *B* lags behind that of group A_{All} at any point on the scale. Candidates who select social sciences (SOC) are completely inferior in ranks with regard to all Matura candidates (Zupanc & Bren, 2009).

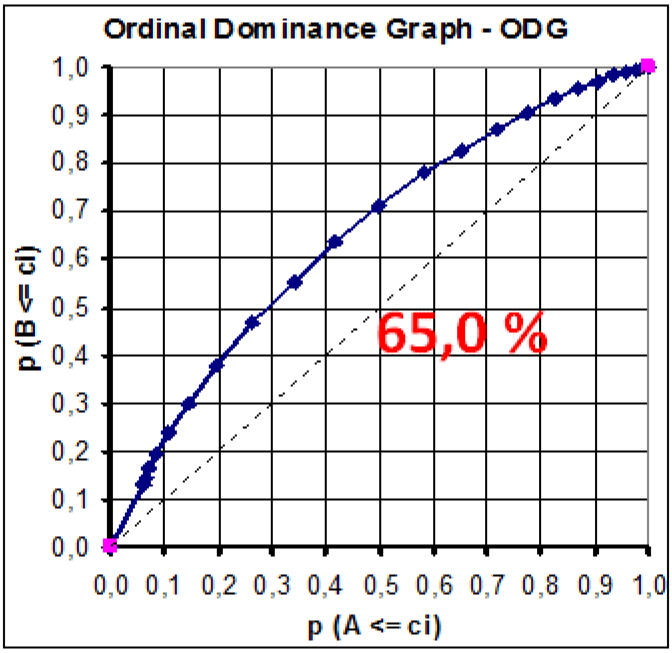
http://www.iaea.info/documents/paper_4d73f71.PDF.

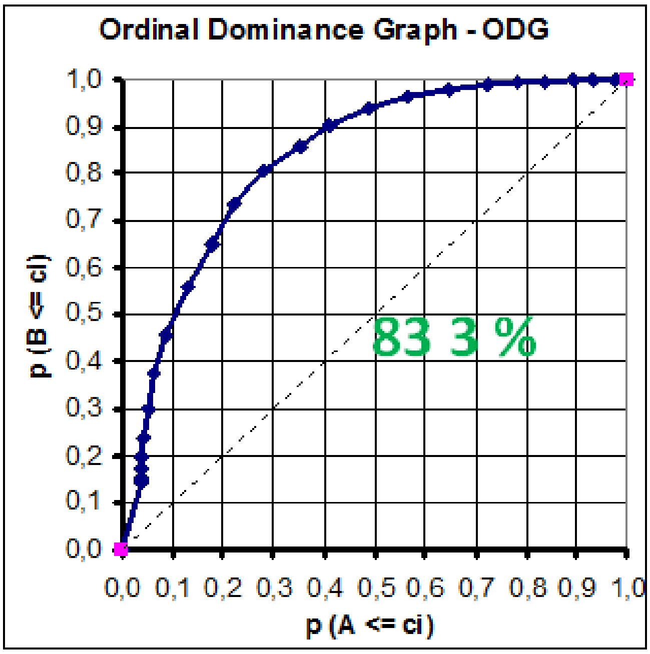
Comparison of achievement distributions in three compulsory subjects at *Matura* 2009 in Slovenia for the candidates **with at least one optional natural-science subject** vs. candidates with no optional natural-science subject (Left graph below)

Table and left graph below: Approximation of the MWW *U* test *z* value, stochastic superiority *ρ* and directed dissimilarity *d* for Matura 2009 results in the three compulsory subjects for candidates A with at least one optional natural-sciences subject.

Comparison of achievement distributions in three compulsory subjects at *Matura 2009* in Slovenia for candidates **with two optional natural-science subjects** vs. candidates with two optional social-sciences subjects (Right graph below).

Table and right graph below: Approximation of the MWW U test z value, stochastic superiority *ρ* and directed dissimilarity *d* for Matura 2009 results in three compulsory subjects for candidates A with two optional natural-science subjects (N_A = 588) vs. candidates B with two optional social-science subjects (N_B = 1,919).





AtLeast1Sci / No_Sci	*N*	*U*	*z*	*p*	*ρ*	*d*	Difference
A AtLeast1Sci	2499	11866599	22.483	0.0000	0.650	0.300	Medium
B No_Sci	7305	6388597	-22.483	0.0000	0.350	-0.300	

2x_Sci / 2x_SocSci	*N*	*U*	*z*	*p*	*ρ*	*d*	Difference
A 2x_Sci	588	939483	24.498	0.0000	0.833	0.665	Large
B 2x_Soc Sci	1919	188890	-24.498	0.0000	0.167	-0.665	

What is evident from the data is that the majority – nearly three times more students – select social-science optional subjects (*Matura* exam 2009): at least one optional natural-science subject = 2,499 vs. no optional natural-science subject = 7,305, and two optional natural-science subjects = 588 vs. two social-science optional subjects = 1,919.
Possible causes for this could be that the majority of students avoid traditionally exacting natural-science subjects with the aim od selecting easier way to pass the *Matura* exam. In this way, a consensus should be obtained between the natural and social sciences community not to drop educational standards and to give students clear feedback on their knowledge, differencing which knowledge stage is sufficient and which is not, which is good, and which is excellent (Zupanc, Bren, 2010).
The aforementioned does not apply to all optional *Matura* subjects in Slovenia; philosophy, as a humanistic discipline, has similar characteristics like science *Matura* subjects. Philosophy is chosen by a minority, candidates are good overall achievers on average, some of them were rewarded in the International Philosophy Olympiad (IPO)
In Slovenia, the admission procedure for enrolment into the first year of tertiary education is centralized through the Admissions Office. Candidates set their preferences and the selection process is performed based on their USE and *Matura* exam achievements, as well as on enrolment capacity of higher education.
By law, faculties can influence the selection process only through the enrolment priority formula by placing adequate weights on subjects fundamental to their study.
The Faculty of Medicine already weights *Matura* achievements of its candidates in Chemistry, Biology and Physics, so that the students enrolled would have satisfactory fundamental knowledge ensuring a proper building block for successful studies.
Through data analysis, other faculties could follow this best practice as well.

Zupanc, D. & Bren, M. (2010). Izbira predmetov pri maturi in splošni uspeh - ali naravoslovne predmete izberejo po uspehu boljši? *Šolsko polje, 21(3/4)*, pp. 107-135.

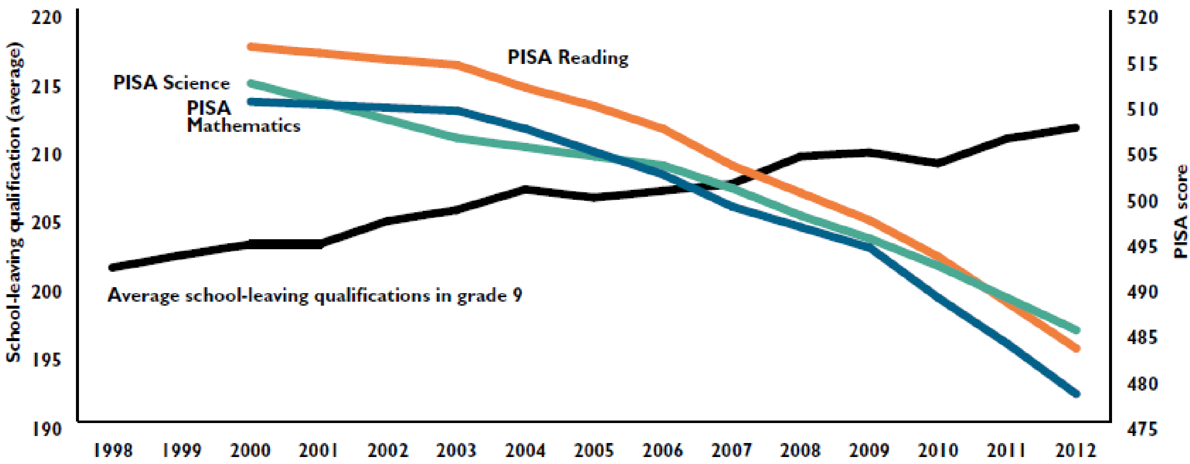
2. Internal and external grading

In some countries, internal grading does not reflect achieved educational standards. If an upward shift of average school grades over an extended period of time happens without a corresponding increase in student achievement, we refer to this as inflation of grades. This issue was pointed out years ago when experts in the U.S. became aware of the year-to-year rise in the grade point average (GPA), and also when a similar situation occurred with the internal grading of various subjects of general *Matura* in Slovenia.

Zupanc, D. & Bren, M. (2010). Inflacija pri internem ocenjevanju v Sloveniji. Sodobna pedagogika, 61(3), pp. 208-228.

Sweden has a big problem with inflation of internal grades.
In Figure 9, Sweden's international educational performance is compared with the average school-leaving qualifications gained by grade 9 pupils, i.e. the year group which is tested by PISA. The curve, which shows average school-leaving grades, is basically a reverse mirror image of how Sweden performed in the PISA study. When the Swedish PISA scores were highest (when the first tests were conducted in 2000), the average school-leaving grade was around 201 points (maximum 320). The average school-leaving grade rises almost continuously year per year, while the opposite applies to the PISA scores. The average school-leaving grade increases by about 5 percent (corresponding to about 0.15 standard deviation units), while PISA scores fall by about 6 percent. This is perhaps the most powerful evidence we have of grade inflation in Swedish compulsory schools in the 2000s. (Henrekson, & Jävervall, 2016: 16)

Figure 9: Sweden's average PISA scores (2000–2012) and the average school-leaving qualifications in grade 9 (1998–2012).



Source: Swedish National Agency for Education (2001, 2004a, 2007, 2010, 2013), SIRIS (Swedish National Agency for Education) and OECD (2015).

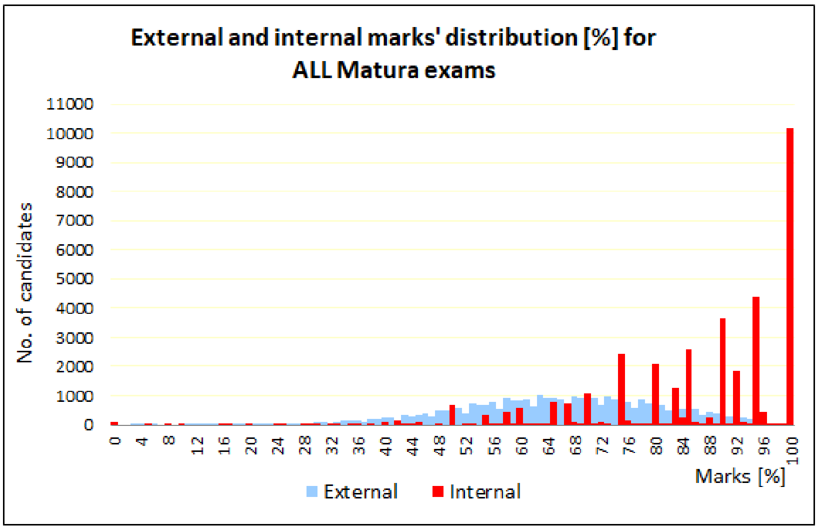
Henrekson, M. & Jävervall, S. (2016). p. 16

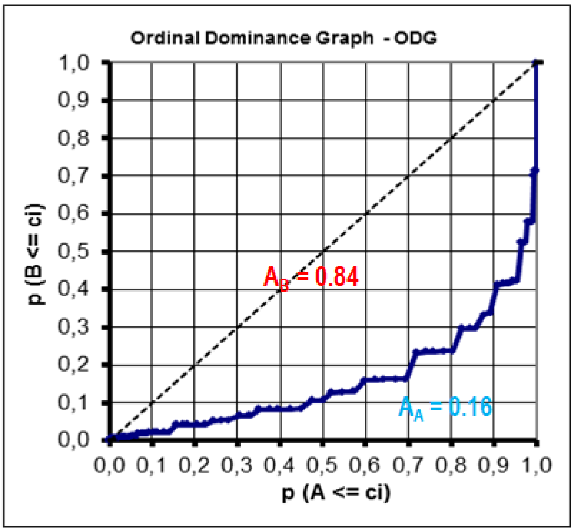
Internal and external grading in Slovene Matura exams

Matura grades in Slovenia are composed **of External (blue - 80% weight)** and **Internal (red – 20% weight)** marks

The distributions are completely different; in almost 1/3 of the internally graded exams the score is 100 % (N = 36,000 in Year 2011)

The probability that a randomly chosen candidate will have a higher internal than external mark (when both are recalculated to 100%) is 84% (N= 36,000 in 2011)





Matura marks [%]	*N*	*U*	*z*	*p*	*ρ*	*d*	Difference
External	35791	200944186	-159,073	0,0000000	0,157	-0,686	Large
Internal	35717	1077402961	159,073	0,0000000	0,843	0,686	

Internal grades in Slovenia are inflated (Zupanc, Bren, 2010), meaning that higher grades do not derive from a corresponding increase in student achievement. That is confirmed by correlational research between the internal and the external part of the Matura exams (correlations are relatively low) and by the Matura Subject Commission's trial, where teachers' and external examiners' grades regarding the same objectives were compared.

For Slovenia, Zupanc and Bren's study (2010) confirms the prognosis given nine years before, which predicted an annual increase in internal grades in the school-based parts of the *Matura* exams (i.e., oral exam, laboratory work, coursework, etc.); these are now approaching a 90% average, and even exceed that percentage in some subjects.

The National Examinations Centre (NEC) in Slovenia has 3 different measurements of knowledge for all candidates in each subject at the end of Upper Secondary Education (USE) – Gimnazija programme

EXTERNAL MARKS in [%] (externally graded paper and pencil exams) – BLUE full line ___________

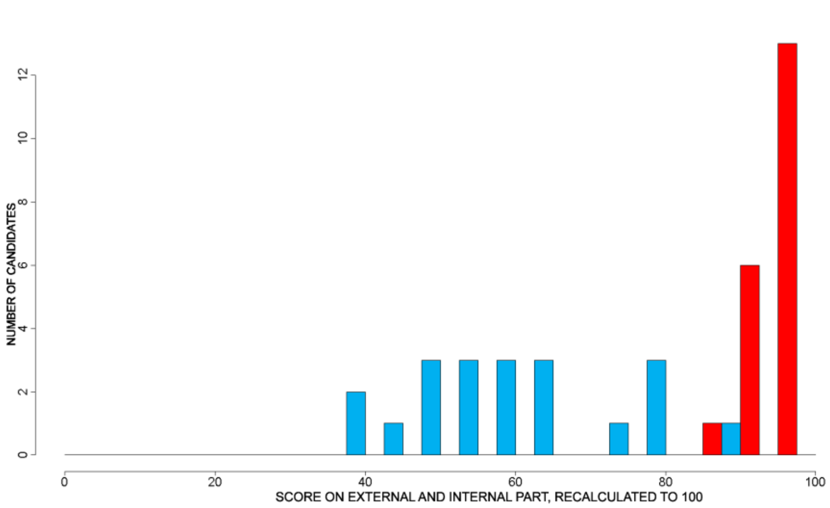
INTERNAL MARKS in [%] (internally graded Matura subjects' coursework, oral exams…) – RED dotted line …………..

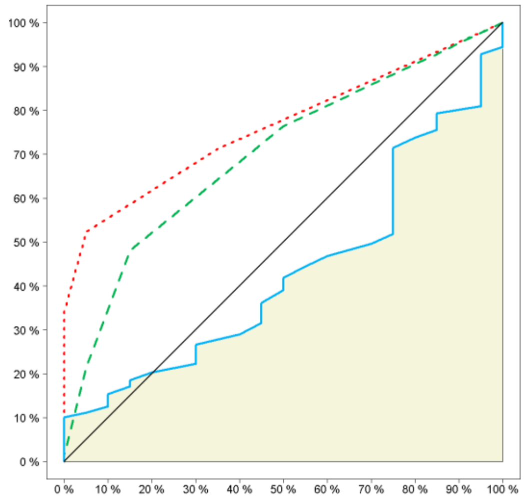
TEACHERS' INTERNAL GRADES from 1 to 5 in the last year of study (teachers' grades before Matura exams -) – GREEN dashed line ------------------

We will compare each of the three different distributions with the corresponding national distribution! There are big differences in these three distributions between schools (Gimnazija)!

With Ordinal Dominance Graphs (ODG) (www.iaea.info/documents/paper_4d73f71.PDF), we can present multiple comparison pairs of distributions in one picture. Even if each pair of distribution has different scale: grades from 1 to 5, percentage marks, raw scores etc.

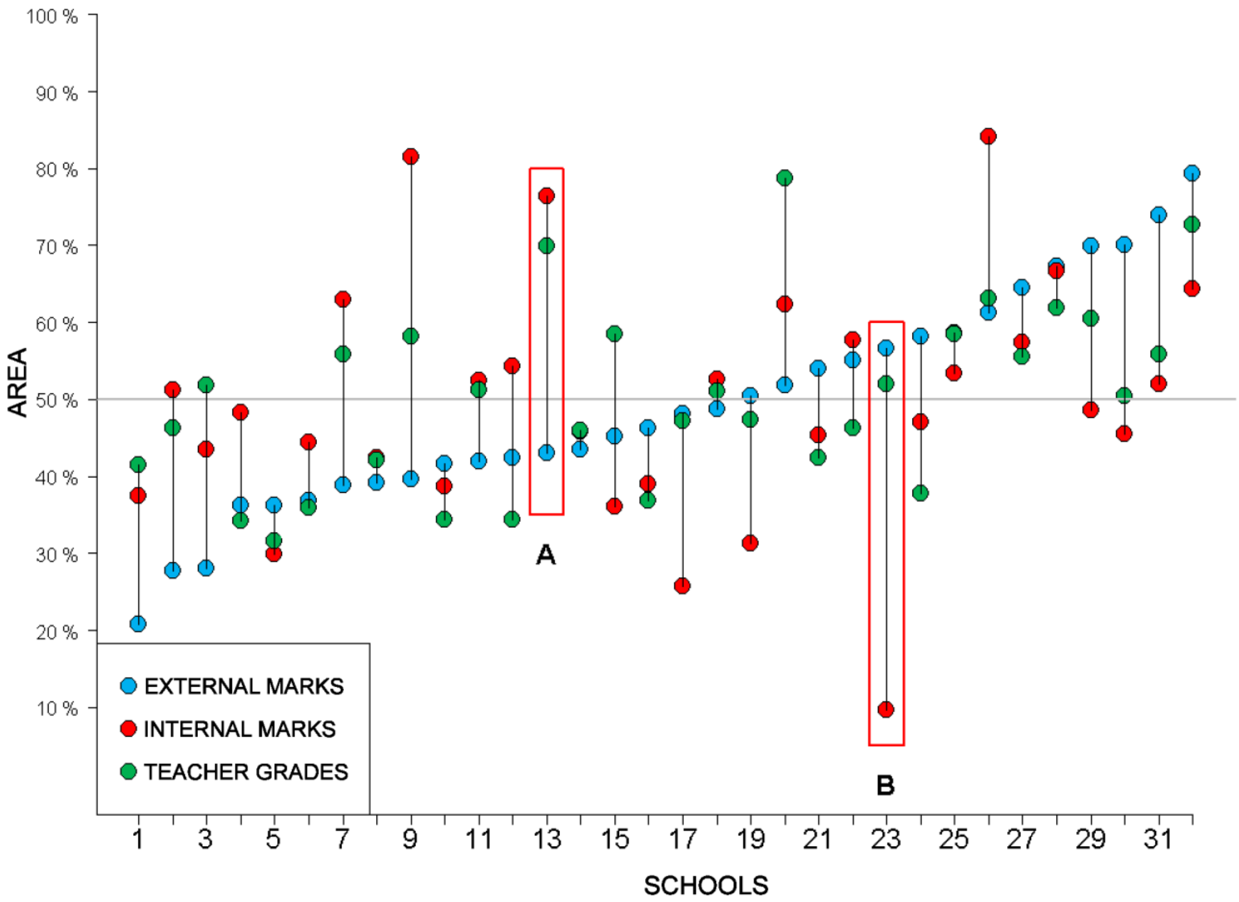
External and Internal marks' distributions [%] and ODGs for 3 different pairs of achievement distributions (Spring Matura session 2011, School A, Physics, N = 20)





EXTERNAL MARKS – BLUE full line INTERNAL MARKS RED dotted line TEACHERS' GRADES GREEN dashed line

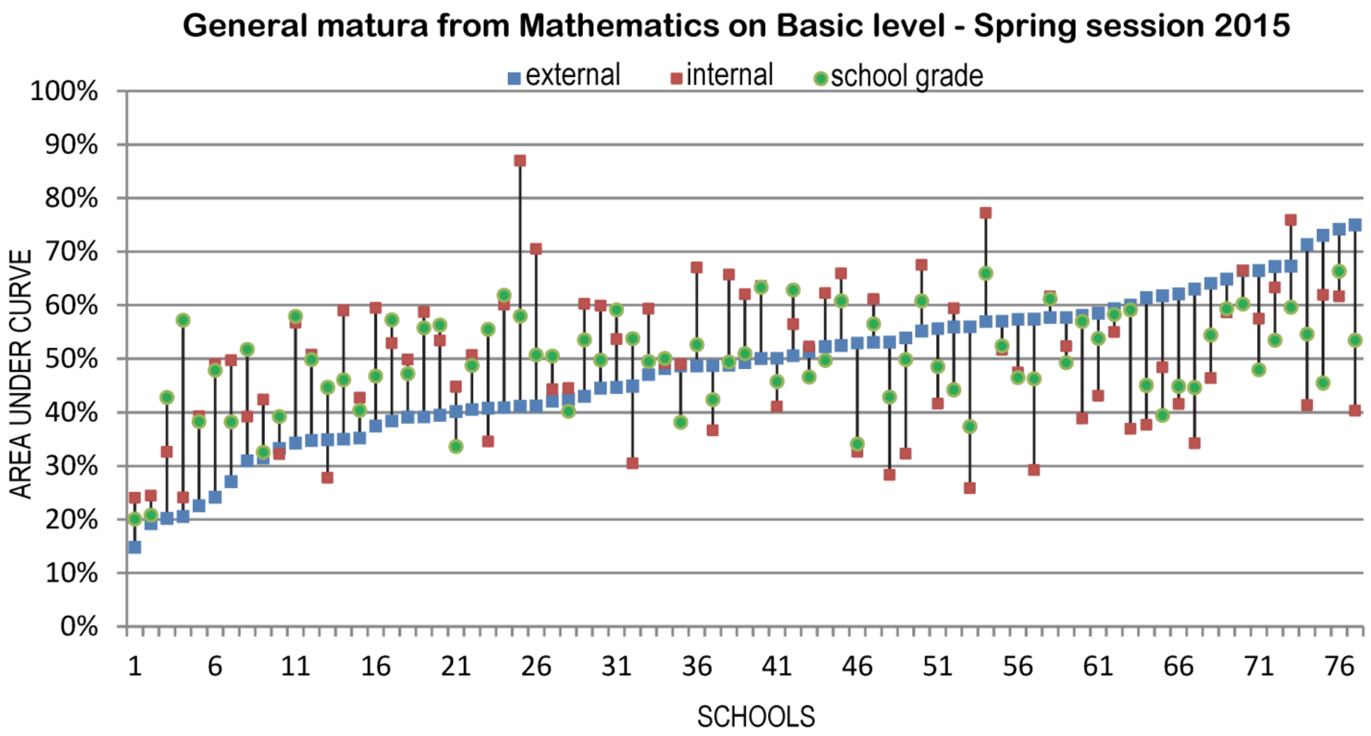
Comparing areas from ODGs for three different pairs of achievement distributions (Spring *Matura* session 2011, Physics, 32 Schools, N_{Sch} >= 20)



Internal grading, performed by candidates' teachers, is not subject to absolute criterion; teachers give the spectrum of lower and higher grades **regardless** of measured achievements **relative to absolute (criterion) reference.**
As a rule, high achievements for the external parts is followed by low achievements for the internal parts and vice-versa; low achievements for the external parts is increased by high achievements for the internal parts.

The pattern is even more obvious comparing areas from ODGs for three different pairs of achievement distributions for Mature candidates (Spring session 2015) in Mathematics on Basic level in 84 Schools (Kokol, 2016) – Graph below.

Very different distributions for external and internal grading in schools are disturbing and iniquitous; we have problems with equity!
An issue arises, if two candidates get different grades for the same achievement; however, it is also problematic, if two candidates get the same grade for very different achievements.



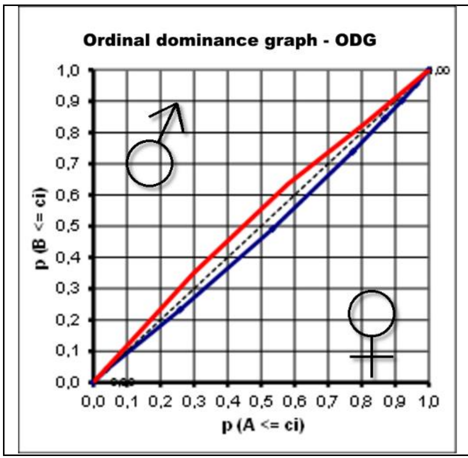
KOKOL, S. (2016).. *Uravnoteženost ocen na splošni maturi : magistrsko delo*. Ormož: 127 f..

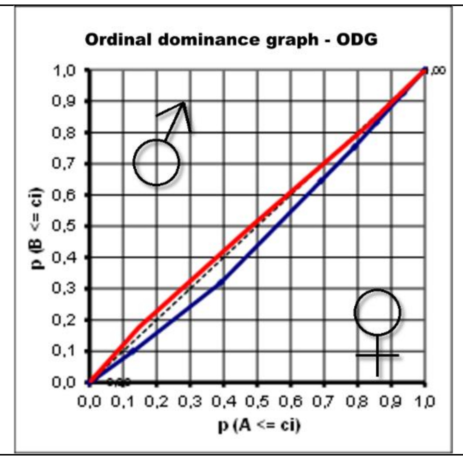
https://dk.um.si/Iskanje.php?type=enostavno&niz=Uravnote%C5%BEenost%20ocen%20na%20splo%C5%A1ni%20maturi&vir=dk&lang=slv.

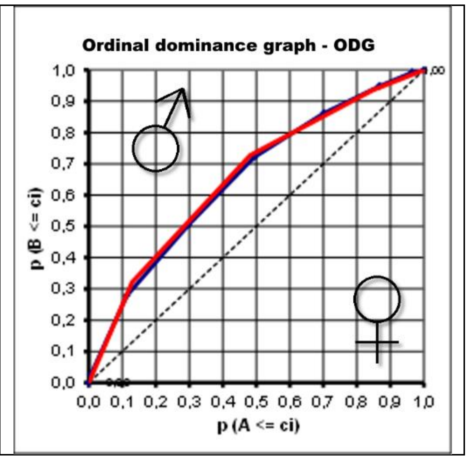
The criterion is higher for better classes and lower for classes with lower achievement students; **according to** literature, internal grading is biased with regard to Socio-Economic Status (SES), gender, etc.
According to the OECD (2012) study, teachers with internal grading tend to give advantaged SES students and girls better school grades, even if they do not have better performance and attitudes than disadvantaged SES students and boys.
School grades are positively and significantly associated with students' expectations of completing a university degree, even after accounting for students' performance and their SES (OECD, 2012).
When comparing students of similar SES and academic achievement, students who expect to graduate from university are more likely to complete these degrees than their peers who do not hold such high expectations (Campbell, 1983; Carbonaro et al., 2011; Morgan, 2005; OECD, 2012; Perna, 2000; Sewell et al., 2003).
Differences in expectations between students with similar performance levels who have either advantaged or disadvantaged SES are highest in Korea, Latvia, Poland, Serbia and Slovenia. In all these countries (OECD, 2012),

disadvantaged SES students are four times less likely to expect to graduate from university than advantaged students who perform equally well.
Differences in expectations and school grades are also the origins of differentiation in enrolment into vocational or academic programmes.
Among boys and girls who perform equally well in PISA reading and mathematics, girls are more likely to expect to complete a university degree. Boys are more likely than girls to expect to end their formal education in upper secondary school, even if they do just as well as girls on the PISA assessment (OECD, 2012)
Inequalities in internal grading lead to inequalities in educational expectations, and, later, to inequalities in educational attainment and labor-market outcomes, thus perpetuating social disparities and reducing opportunities for upward mobility, especially among disadvantaged boys.

Slovenia case: External assessment NARROWS the gap between boys and girls (general Matura 2015)







Mathematics **English** **Mother Tongue**

Comparing distribution of grades between final year in upper secondary school and Matura examination – **Mathematics** at the LEFT (red – internal school grades, blue – Matura grades).

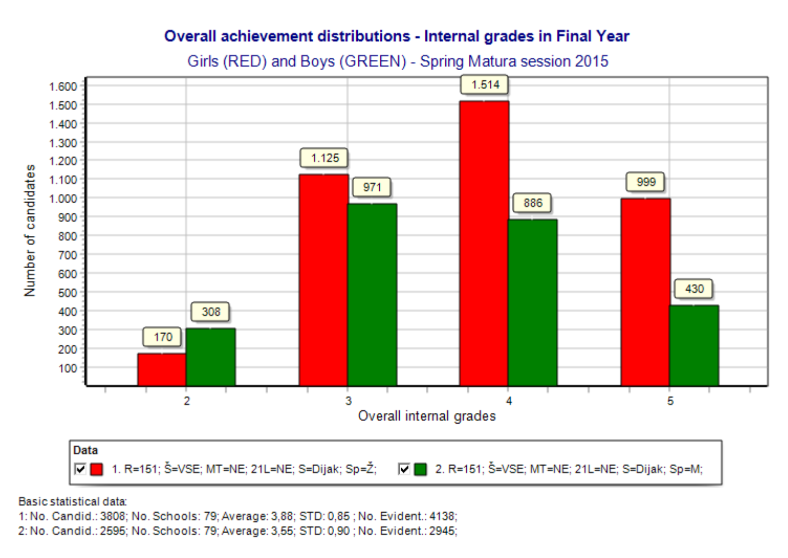
Comparing distribution of grades between final year in upper secondary school and Matura examination – **English** in the MIDDLE (red – school grades, blue – Matura grades).

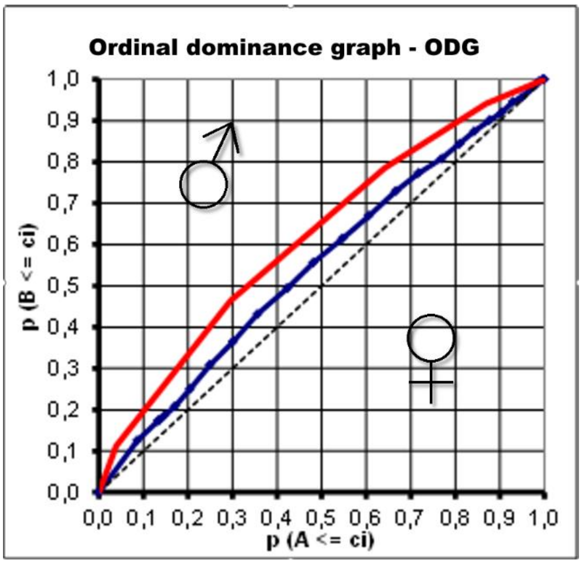
Comparing distribution of grades between final year in upper secondary school and Matura examination – **Mother tongue** at the RIGHT (red – school grades, blue – Matura grades).

If we observe differences between grades boys and girls receive in the final year of upper secondary school and at Matura examination for the same subject (Mathematics or English), external examination (Matura) narrows the gap between boys and girls.

There is a big difference between boys and girls in **Mother tongue - Slovene language**, and this difference is persistent across grades received in the final year of upper secondary school and at Matura examination. Causes underlying this difference go back to education in primary and lower secondary schooling, as effects can be seen in students' results at the end of compulsory schooling (at the age of 14). The results can also be seen in PISA reading literacy. Both results suggest this difference is not exclusively about internal/external grading.

In summary for overall achievement, we observe differences between internal grades boys and girls receive in the final year of upper secondary school and at Matura examination. External examination (Matura) narrows the gap between boys and girls.



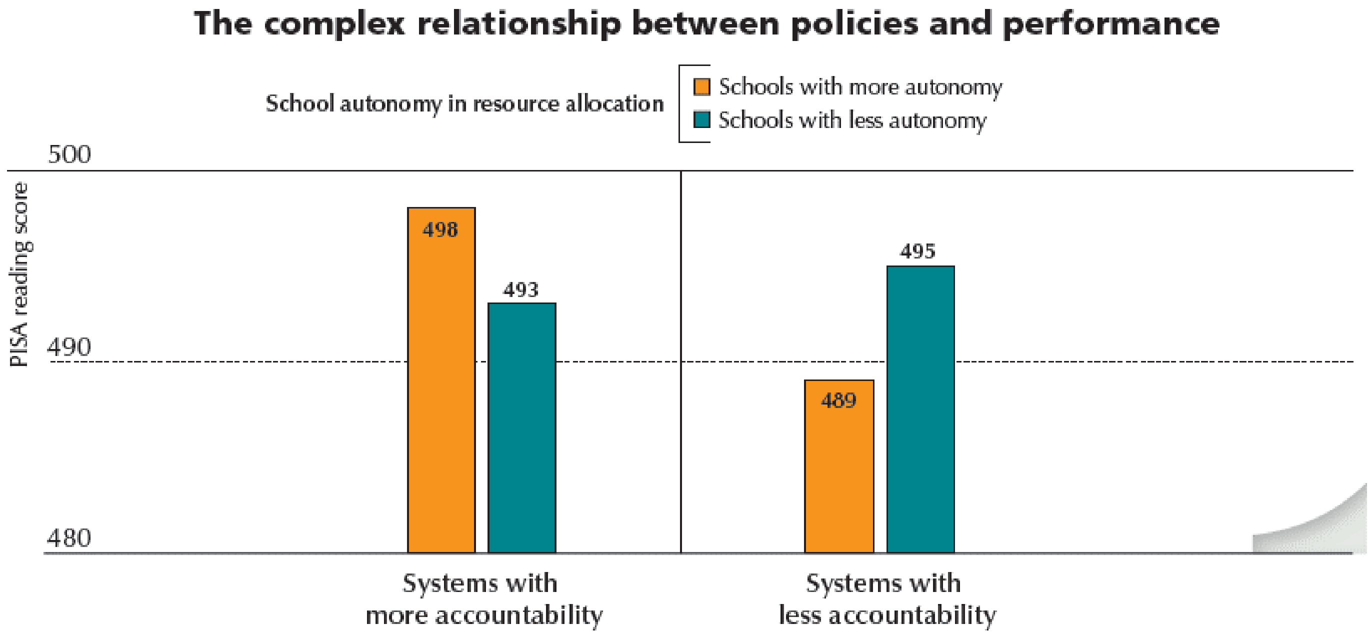


VIII. School autonomy and accountability

Speaking about school autonomy within OECD educational studies, we refer to **autonomy in curricula and assessments** (establishing student-assessment policies, choosing which textbooks are used, determining course content, deciding which courses are offered), as well as to **autonomy in resource allocation** (responsibility for selecting teachers, for hiring and dismissing teachers, establishing teachers' starting salaries, determining teachers' salary raises, formulating the school budget, and deciding on budget allocations within the school) (OECD, 2011).

One possibility of **accountability** within the country is Curriculum-Based External Examination Systems (CBEESs) (OECD 2010, p 46). Another possibility is that schools are more **accountable** to students, parents and the public at large for their outcomes, if achievement data are made available to the public (OECD, 2011).

The relationship between autonomy, accountability and performance is complex. It is a combination of several autonomy and accountability policies, not just a single, isolated policy relating to better student outcomes (OECD, 2011).



OECD (2011, October). School autonomy and accountability: Are they related to student performance? *PISA IN FOCUS 2011/9.*

Across OECD countries, those that use Curriculum-Based External Examination Systems (CBEES) tend to perform higher, even when accounting for national income: students in school systems that use CBEES perform, on average across OECD countries, 16 points higher than students in school systems that do not use these examinations (OECD 2010, p 46).

In countries where schools account for their results by posting achievement data publicly, schools that enjoy greater autonomy in resource allocation tend to show marginally better student performance than those with less autonomy.

In countries without such accountability arrangements, schools with greater autonomy in resource allocation tend to perform worse (OECD, 2011).

There is no clear relationship between autonomy in resource allocation and performance at country level. PISA results suggest that, when autonomy and accountability are combined in an intelligent manner, they tend to be associated with better student performance (OECD, 2011).

School's autonomy over curriculum & School's autonomy over teacher salaries

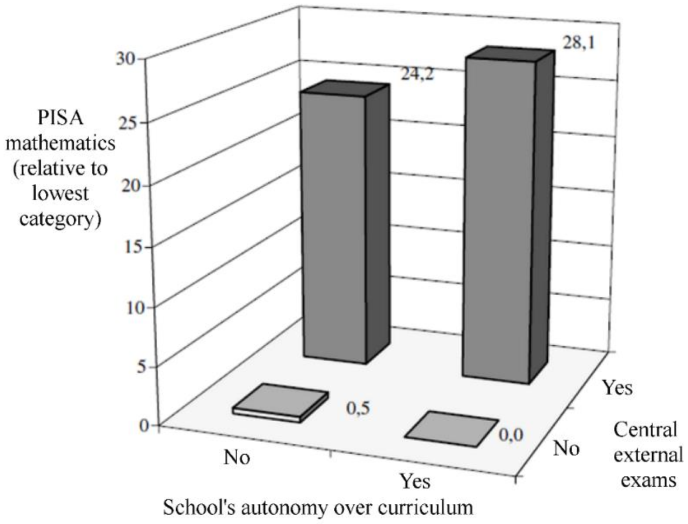
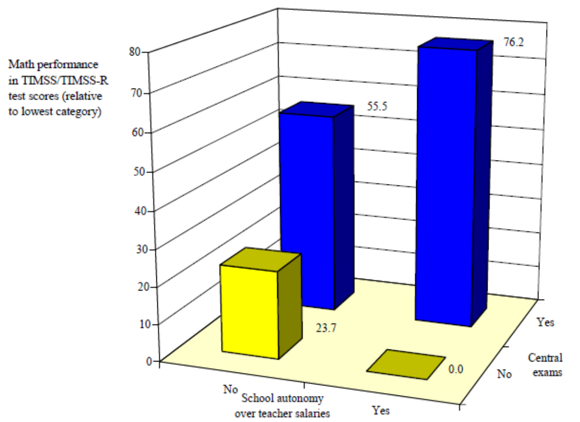


Figure 6: External exams, school autonomy, and student achievement across countries



Notes: Performance difference between the four categories relative to the lowest category which is set equal to zero. Based on a cross-country student-level multiple regression using the combined TIMSS and TIMSS-Repeat micro databases that extensively controls for family background, school inputs, and other institutional features.
Source: Woessmann (2005c).

Wößmann, L. (2008). Zentrale Abschlussprüfungen und Schülerleistungen. Individualanalysen anhand von vier internationalen Tests. Zeitschrift für Pädagogik 54 (2008) 6, S. 810-826 LEFT Picture

Woessmann, L. (2005). The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics, 13* (2), 143-169.

Hanushek, E. A., & Ludger Woessmann (2011). The Economics of International Differences in Educational Achievement. In: Eric A. Hanushek, Stephen Machin and Ludger Woessmann (ed.). *Handbook of the Economics of Education, Volume 3*. Amsterdam: North Holland, pp. 89-200. Figure 6: RIGHT Picture

Autonomy and accountability go together

Greater autonomy in decisions relating to curricula, assessments and resource allocation tends to be associated with better student performance, particularly when schools operate within a culture of accountability.

There is an indication that local decision-making (autonomy) works better when there is also external accountability that limits any opportunistic behaviour of schools.

School autonomy in subject curriculum has no positive effect on achievement, unless there are central external exams (CBEES) in the country.

School autonomy over teacher salaries has no positive effect on achievement, unless there are central external exams (CBEES) in the country or economy.

School autonomy has positive effects in developed countries and where external exit exams introduce accountability, but negative effects in developing countries with external exit exams (Wößmann, 2016).

Does school autonomy make sense everywhere? Our results suggest that the answer is a clear "no" (Hanushek, Link, Woessmann, 2013).)

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA. *Journal of Development Economics,* 104, pp. 212-232.
Wößmann, L. (2016). "The Importance of School Systems: Evidence from International Differences in Student Achievement", *Journal of Economic Perspectives 30 (3),* 03-31.

IX. Equity in education

It is frequently assumed that effectiveness and equity objectives are mutually exclusive. It is too often the case that existing education and training systems reproduce or even compound existing inequities (COM; 2006).

Equity "is viewed as the extent to which individuals can take advantage of education and training, in terms of opportunities, access, treatment and outcomes. **Equitable systems ensure that also the outcomes of education and training are independent of socio-economic background** (SES) and other factors that lead to educational disadvantage, and that treatment reflects individuals' specific learning needs" (EU, 2006).

All countries should try **to meet two goals** to ensure the best outcome for their schools: **getting high levels of student achievement** while **minimizing systematic gaps in performance** (Wößmann, 2016).

Students rely on school marks. In all countries and economies that disseminated the Educational Career Questionnaire (ECQ) in the PISA 2009 cycle and were asked about the marks they received, **students with higher marks were more likely to expect to graduate from university than students with lower marks. Marks are relevant in predicting students' expectations of graduating from university,** even after accounting for student performance and ISCED programme (OECD, 2012).

This means that students rely on information about their performance that receive in school to form their long-term expectations; and they should: research shows that marks are related to students' long-term outcomes, like earnings and university completion (Rosenbaum, 2001). Students do well to take this information into account.
Teachers reward different sets of factors through marks, which, in some cases, have little to do with students' mastery of skills or attitudes and behaviours that foster learning. Teachers and schools would do well to understand that **marks are more than an appraisal of skills and knowledge acquired; they provide information to the student about what is valued not only in school but in life in general** (OECD, 2012; 65).

Rosenbaum, J. (2001), Beyond College for All: Career Paths for the Forgotten Half, Russell Sage Foundation, New York, New York.

The differences in students' expectations of completing a university degree that are related to socio-economic status are not substantially explained by differences in the school marks advantaged and disadvantaged students receive in primary and secondary education. After accounting for performance in reading and mathematics, marks account for a small part of the inequality in expectations related to socioeconomic status (SES). The remaining differences are still statistically significant in practically all countries and economies (OECD, 2012).

Differences in expectations between students who have similar performance levels but who come from either advantaged or disadvantaged backgrounds **is highest in** Korea, Latvia, **Serbia** and **Slovenia**. In all these countries socio-economically advantaged students **are four times less likely to expect to end their education after completing upper secondary school than disadvantaged students who perform equally well** (OECD, 2012, 35-36). They are also large in **Croatia, Hungary**, Iceland and Poland, where a one-unit increase on the PISA index of economic, social and cultural background is seen to at least double the likelihood that a student will expect to complete a university degree.

These remaining differences could be due to the influence of peers, parents and teachers, who assume that students from advantaged backgrounds will have higher expectations; differences in access to curricula (academic or vocational) related to socio-economic background; or to the perception held by disadvantaged students that they are less likely to have access to, and succeed in, university, even if they have similar marks and performance levels, and follow the same academic curriculum as their advantaged peers. (OECD, 2012; 65)

To the extent that **such inequalities in expectations constitute a barrier** to eventual enrolment and graduation, they pose **a serious challenge for countries because of the loss of human capital and skills potential they represent**.

Poor children have a much higher risk of low school achievement, but this fact **does not mean that increasing the income of poor parents would automatically improve children's achievement**. (OECD, 2012; 92-93)

Whatever the reason, inequalities in marking practices **may lead to inequalities in** educational expectations, and, later, to inequalities in educational attainment and labour-market outcomes, thus **perpetuating social disparities and reducing opportunities for upward mobility,** especially among disadvantaged boys (OECD, 2012; 70).

OECD (2012), Grade Expectations: How Marks and Education Policies Shape Students' Ambitions, PISA, OECD Publishing.
ISBN 978-92-64-18750-4 (print) ISBN 978-92-64-18752-8 (PDF)
EU (2006). Efficiency and equity in European education and training systems. Summaries of EU legislation: http://eur-lex.europa.eu/legal-content/GA/TXT/?uri=URISERV:c11095
COM (2006). 0481 final. Communication from the Commission to the Council AND to the European Parliament, Efficiency and equity in European education and training systems {SEC(2006) 1096}, Brussels, 8.9.2006

1. Early tracking and equity

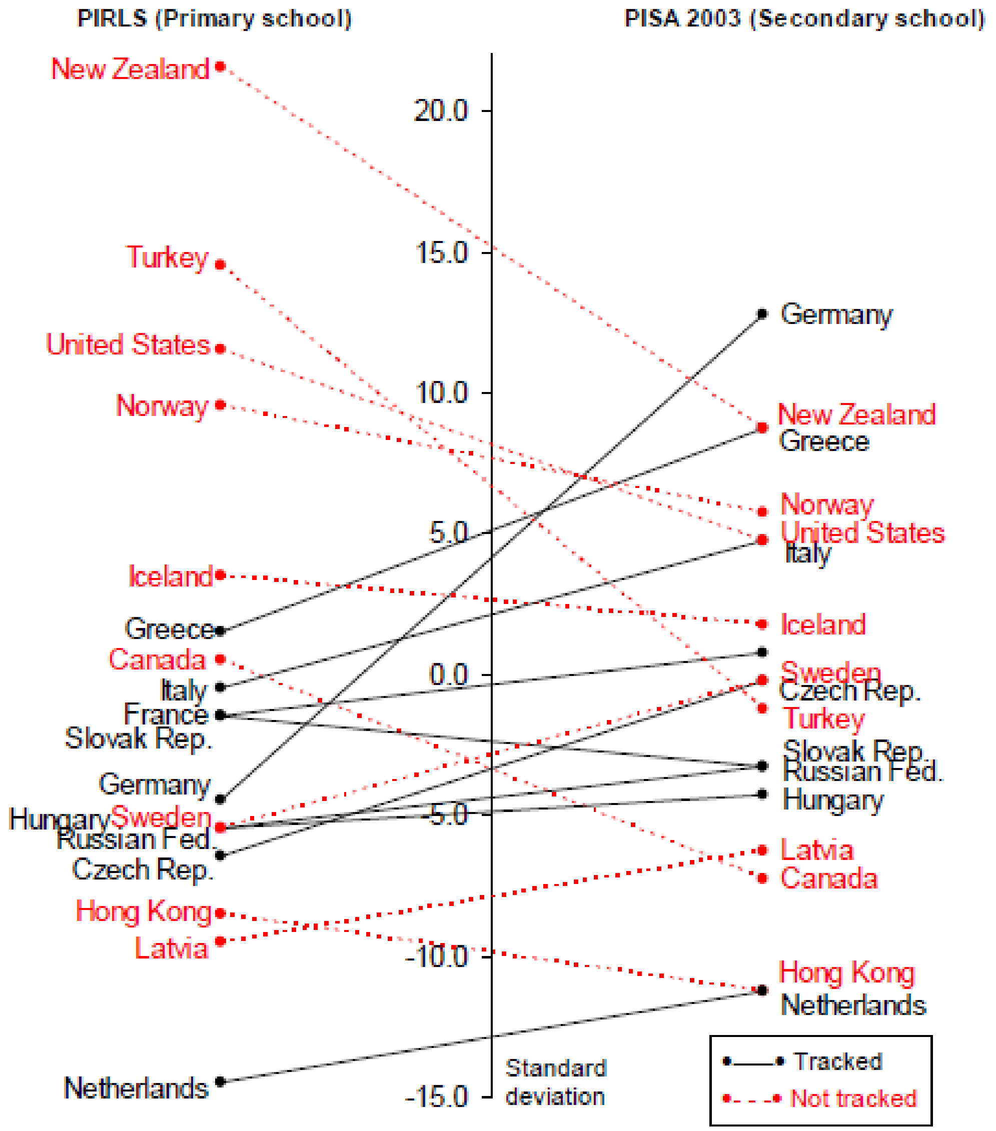
Most countries find that the performance of students varies systematically with a variety of characteristics. The most prominent **concerns are generally related to family background**: income, wealth, parents' educational attainment and race, ethnicity.
Early tracking into differing-ability schools seems to increase inequality in achievement without increasing achievement levels (Wößmann, 2016).

Figure 8 shows the inequality in reading achievement in 4th grade (in PIRLS) and at age 15 (in PISA 2003) for all countries that participated in both studies, measuring educational inequality by the standard deviation in student test scores (Hanushek, Woessmann, 2011). The essence of the analysis is to compare the change in inequality that occurs from primary to lower-secondary school between countries with and without educational tracking during this period.

Figure 8 shows (Hanushek, Woessmann, 2011) that in the majority of cases dispersion of student achievement from primary to lower-secondary school within the country becomes larger, if school systems track their students into different school types before the age of 16. The opposite applies as well; the dispersion of student achievement in becomes smaller the majority of cases, if school systems do not track their students until the age of 16.

Wößmann, L. (2016). "The Importance of School Systems: Evidence from International Differences in Student Achievement", *Journal of Economic Perspectives 30 (3),* 03-31.

Figure 8: Educational inequality in primary and secondary school



Hanushek, E. A., & Woessmann, L. (2011). The Economics of International Differences in Educational Achievement. In: E. A. Hanushek, S. Machin & L. Woessmann (ed.). *Handbook of the Economics of Education, Volume 3*. Amsterdam: North Holland, pp. 89-200.

Description of Figure 8: Dispersion of student achievement measured as standard deviation of test scores in primary school (PIRLS) in comparison with lower-secondary school (PISA 2003), in both cases measured as the difference to the international mean of national standard deviations in each test (Hanushek, Woessmann, 2011).
When looking at the change between the achievement dispersion in PIRLS and PISA, the part of the inequality measured at the end of lower-secondary school that already existed in 4th grade is eliminated. The change is indicated by the lines that connect the two points of each country.

The lines indicate the change in performance dispersion from primary to lower-secondary school.
For countries with early tracking, **solid connecting lines are used**, while dashed lines indicate countries without early tracking. The **black solid lines** indicate **school systems that track their students into different school types before the age of 16**, while the red dashed lines refer to those school systems that do not track their students by this age.

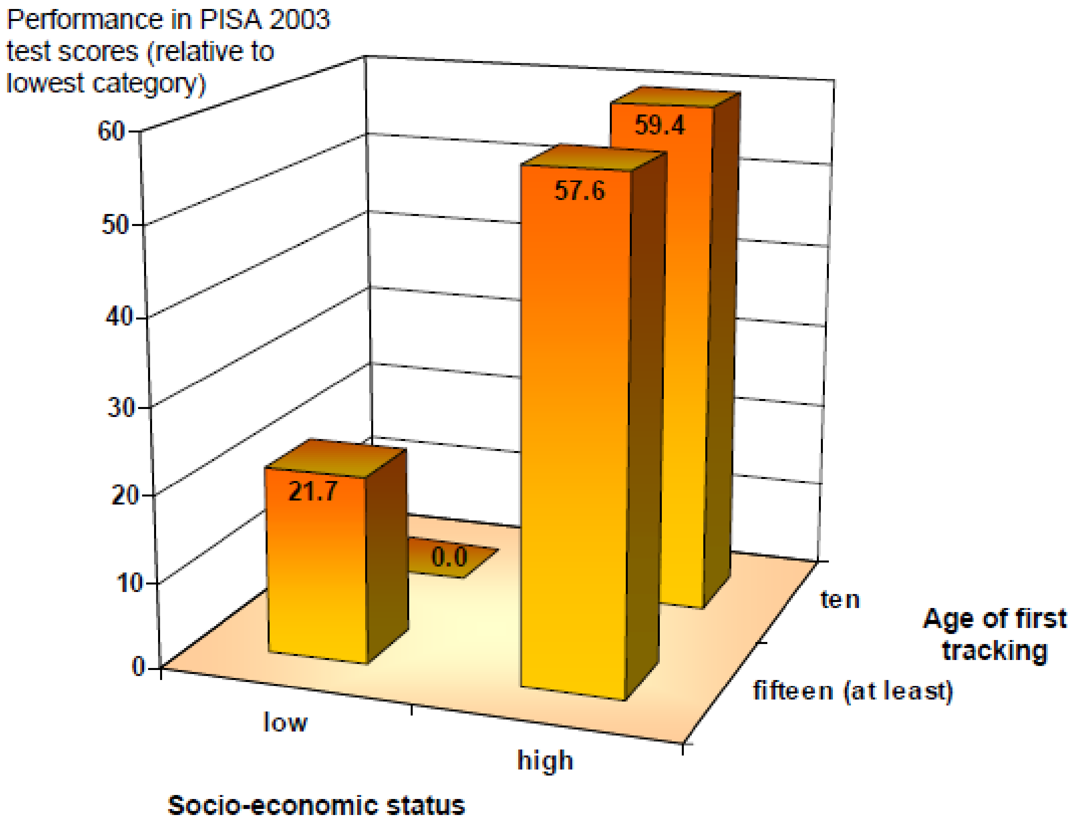
It is evident that nearly all black solid lines point upwards, whereas nearly all red dashed lines point downwards: In countries with early tracking, inequality increases systematically, whereas it decreases in countries without tracking.

Based on a cross-country student-level multiple regression using the PISA 2003 micro database, we can see (Figure 9) that the relation between test scores and family background is significantly smaller the higher the age of first tracking Hanushek, Woessmann, 2011).

As a measure for family background, the PISA study applies the Index of Economic, Social, and Cultural Status (ESCS). Performance differences between the four categories are relative to the lowest category, which is set equal to zero.

In countries with earlier tracking, the achievement difference between children with different socio-economic backgrounds is considerably larger. As the figure reveals, this effect arises primarily from the fact that **children with low socio-economic status in countries without early tracking perform considerably better.**
At the same time, children from families with a relatively high socio-economic status perform at approximately the same level. Accordingly, the effect of later tracking on the average achievement level is again positive, albeit not statistically significant.

Figure 9: Tracking and socio-economic status in PISA



Notes: Performance difference between the four categories relative to the lowest category which is set equal to zero. Based on a cross-country student-level multiple regression using the PISA 2003 micro database that extensively controls for family background, school inputs, and other institutional features. Low and high socio-economic status correspond to the 25th and 75th percentile of the PISA ESCS index, respectively.

Source: Based on Woessmann, Luedemann, Schuetz, and West (2009).

Hanushek, E. A., & Woessmann, L. (2011). The Economics of International Differences in Educational Achievement. In: E. A. Hanushek, S. Machin & L. Woessmann (ed.). *Handbook of the Economics of Education, Volume 3*. Amsterdam: North Holland, pp. 89-200.
Woessmann, L., Luedemann, E., Schuetz, G., & West, M. R. (2009). *School accountability, autonomy, and choice around the world.* Cheltenham, UK: Edward Elgar.

The EU Commission concluded that **those European countries that track pupils at an early age display greater variation in pupil achievement than countries with more integrated school systems** (COM; 2016: 6).

Early tracking has especially negative effects on the achievement levels of disadvantaged children. This is partly because it tends to channel them towards less prestigious forms of education and training.

Postponing tracking until upper secondary level, combined with the possibility to transfer between school types, can reduce segregation and promote equity without diminishing effectiveness.

2. Equity issues with SES are seen after 15 years of age as well (case of Slovenia)

Primary and lower secondary education in Slovenia is organised in a single-structure, nine-year basic school attended by pupils aged 6 to 15. The next stage after nine years of compulsory basic education is two to five year non-compulsory upper secondary education that begins at the age of 15. The system is structured as a single-cycle education programme of two to five years, depending on the programme type. As a rule, students enrol at the age of 15 years. Upper secondary education encompasses (Taštanoska, 2017):

- General education programme, with different types of four-year "Gimnazija" programmes (**Gimnazija, classical Gimnazija**, **technical Gimnazija, Gimnazija of economics, Gimnazija of arts**), which ends with General Matura Exams, and
- Vocational education programme, where students need to pass a programme-ending exam to successfully complete the **short upper secondary vocational** (2years) and **upper secondary vocational education programmes** (3 years), and
- Technical education programme (**upper secondary technical & vocational technical education programmes**), where students need to pass the Vocational Matura to successfully complete the education programme: upper secondary technical education programmes (4 years) and vocational technical education programmes (2 additional years after completing a vocational programme).

Taštanoska, T. (Ed.) (2017). The EDUCATION SYSTEM in the Republic of Slovenia 2016/2017. Ljubljana: The Ministry of Education, Science and Sport of the Republic of Slovenia.

Gimnazija programmes with the General Matura Exam at the end prepare students for continuing studies at higher education university study programmes and integrated postgraduate master's study programmes.

The goal of Upper secondary vocational-technical and technical education with the Vocational Matura at the end is to obtain qualifications for a specific occupation to enter the labour market, and to continue education at the tertiary level as well.

The Gimnazija and General Matura combination is the main upper secondary educational route for entering the most academically demanding and prestigious tertiary education courses.

Disadvantaged students hold lower expectations of further education, even when they perform as well as students from advantaged backgrounds. Inequalities in expectations are present in practically all countries and economies.

As referred to above, differences in expectations between students who have similar performance levels but who come from either advantaged or disadvantaged backgrounds **is highest in** Korea, Latvia, Poland, **Serbia** and **Slovenia**. In all these countries, socio-economically advantaged students **are four times less likely to expect to end their education after completing upper secondary school than disadvantaged students who perform equally well** (OECD, 2012, 35-36)

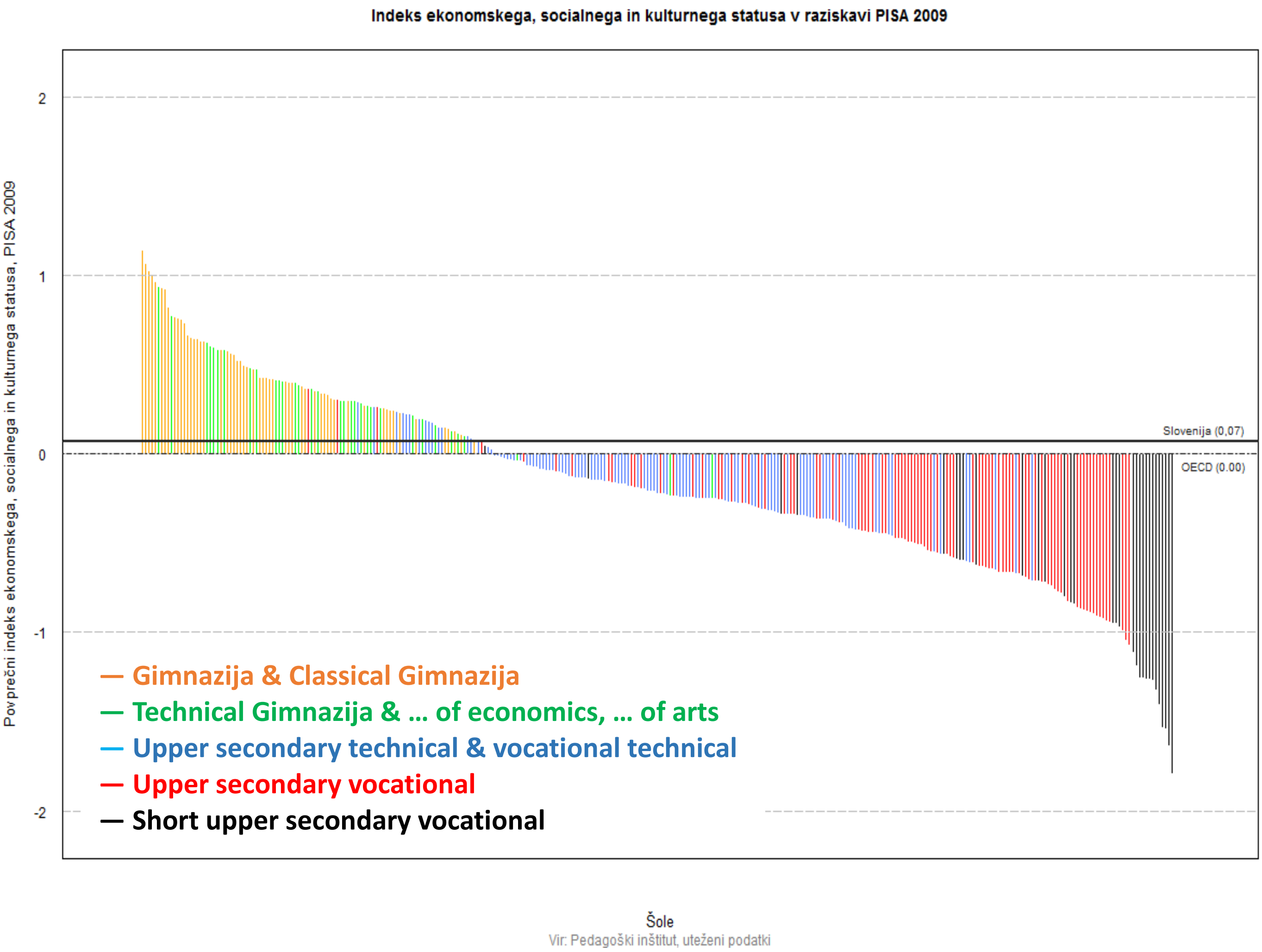
Entering different tracks of upper secondary education in **Slovenia** is highly associated with students and their families' socio-economic status (SES).

OECD (2012), Grade Expectations: How Marks and Education Policies Shape Students' Ambitions, PISA, OECD Publishing.

Secondary school enrollment in **Slovenia** is strongly linked to the students' SES. Students in schools are nested in classes. In the graph below, each colored bar present one class in one upper secondary program in Slovenia in 2009. The height of the bar presents average class SES – Index of Economic, Social, and Cultural Status (ESCS) in PISA 2009.

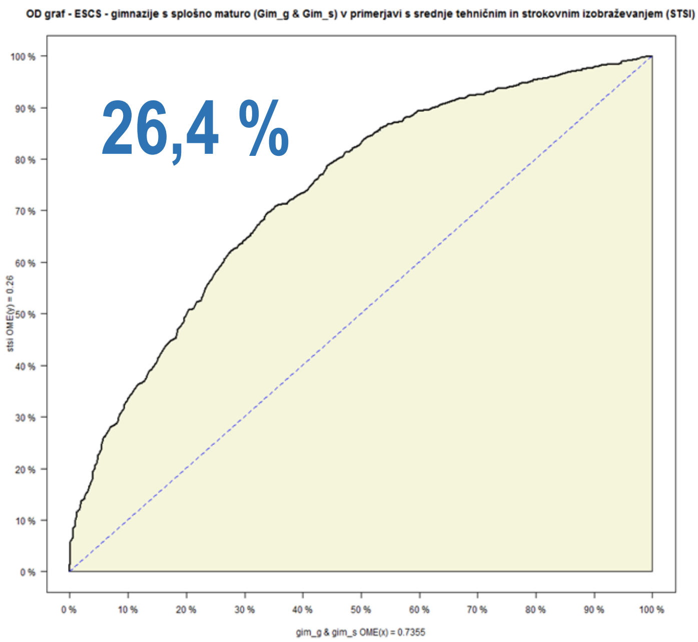
From the PISA data analysis, we can recognize groupings of colored bars in the graph below. Almost all Gimnazija and Classical Gimnazija bars (classes) are on the left and display high above-average SES. The bars of Technical Gimnazija-s, Gimnazija-s of economics and Gimnazija-s of arts are above average on the left as well, along with the rest of Gimnazija bars. Classes from upper secondary technical and vocational technical programmes and corresponding bars are in the middle, with average and below-average class SES. Students with below-average class SES are grouped in 3-year upper secondary vocational programmes on the right side of the graph (below), and students with the lowest SES are grouped in the **2-year short cycle upper secondary vocational**.

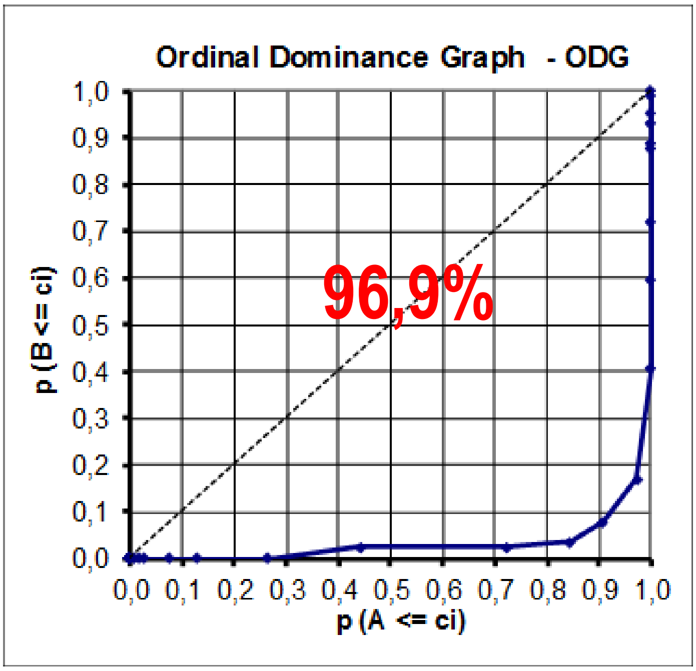
Index of Economic, Social, and Cultural Status (ESCS) in PISA 2009 – for five different upper secondary educational programmes in Slovenia



From the PISA data analysis, we found out that by randomly selecting one student from **Gimnazija program** with **General Matura exam** at the end, and another random student from **upper secondary technical or vocational technical education programmes** in which students need to pass the **Vocational Matura**, and looking at their SES (ESCS)...
… the probability that a **Vocational Matura graduate** will have a higher (or equal) SES (ESCS) than the **General Matura graduate** is only 26.4%.

Students in schools are nested in classes. The probability that an average class SES (ESCS, PISA 2009) in a **Vocational Matura upper secondary school programs** is higher (or equal) than average class SES in the **General Matura upper secondary school programs (Gimnazija)** is only 3.1%.





Composite effect
Being a member of a group (class in school) with a strong socio-economic background has an effect on achievement, which is an additional factor to the effect of the (socio-economic and other) background characteristics of the individual (Van Damme, Opdenakker, Van Landeghem, 2006).
Other researchers have concentrated on the intellectual composition and have made similar findings, identifying positive effects of being a member of a cognitively strong group. A few researchers have established an interaction between individual intelligence and class composition: with regard to achievement, pupils with weaker scores on the intelligence test tended to be more sensitive to class composition.

In research literature, this is often referred to as '**composite effect**'. For example, Opdenakker & van Damme (2001) report that a structure of students in school has important individual and joint effects on student achievement. Identical conclusions are made by Teddlie and Reynolds (2000), who report that the structure of SES significantly influences learning achievement even after they controlled for individual abilities and social standing (Teddlie & Reynolds, 2000, str. 184).

Van Damme, J., Opdenakker, M., Van Landeghem, G. (2006). Educational Effectiveness: An Introduction to international and Flemish research on schools, teachers and classes. Lueven: Uitgeverij Acco. Pp. 28-29

Bornstein, M. H., Hahn, C., Suwalsky, J. T. D., & Haynes, O. M. (2003). Socioeconomic Status, Parenting, and Child Development: The Hollingshead Four-Factor Index of Social Status and the Socioeconomic Index of Occupation. In M. C. Bornstein, & R. H. Bradley, (Eds.). Socioeconomic status, parenting, and child development, pp. 29-82. Mahwah, NJ: Lawrence Erlbaum Associates.

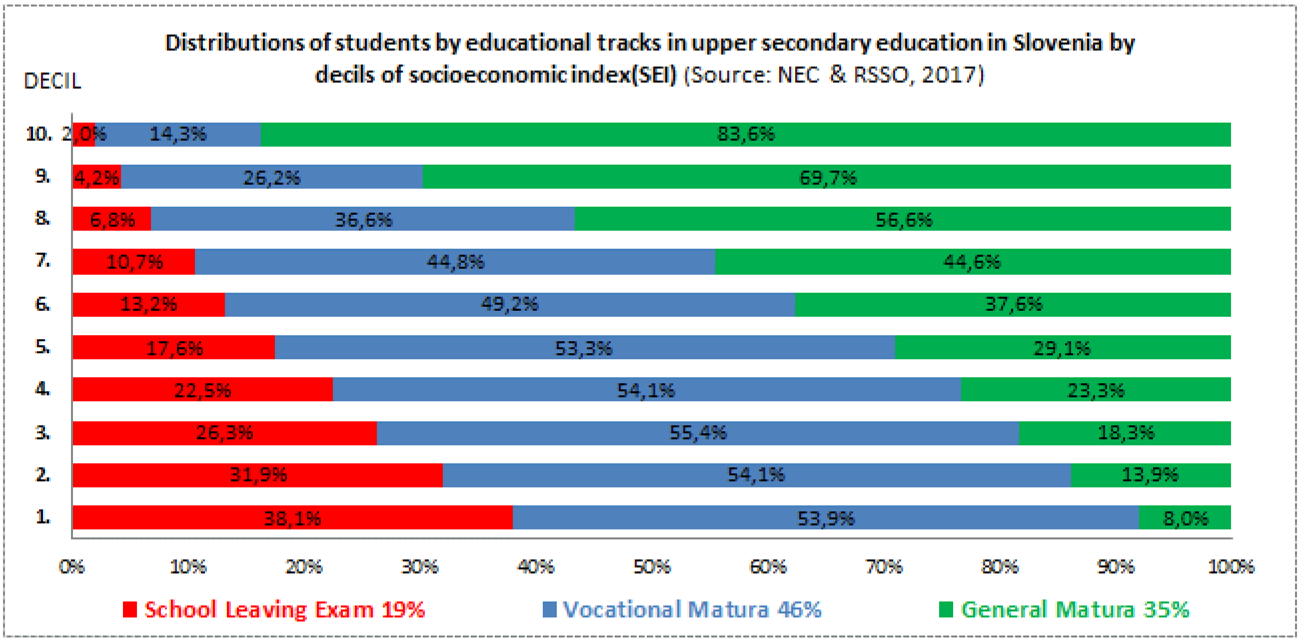
"Equity" in transition from lower secondary to upper secondary school in Slovenia?
The National Examinations Centre (NEC) constructed the newly developed common SES indicator – Slovenian Social Economic Index (SEI) using data from the Republic of Slovenia Statistical Office (RSSO).

The achievements of students from NEC databases originated from 2010-2013, while the administrative database of the RSSO mostly relates to 2011 when the registration census of the population was carried out. By connecting databases, Cankar, Bren and Zupanc (2017) recorded the identities of parents living with the student in the household. They collected data data on education, occupation, income, the value of the property in their possession, etc. for each student's parent (Cankar, Bren, Zupanc, 2017).

Included predictive variables of the socio-economic situation of the family of the student (gross income of parents, value of real estate, education of parents, SES of the profession parents) proved to be important.

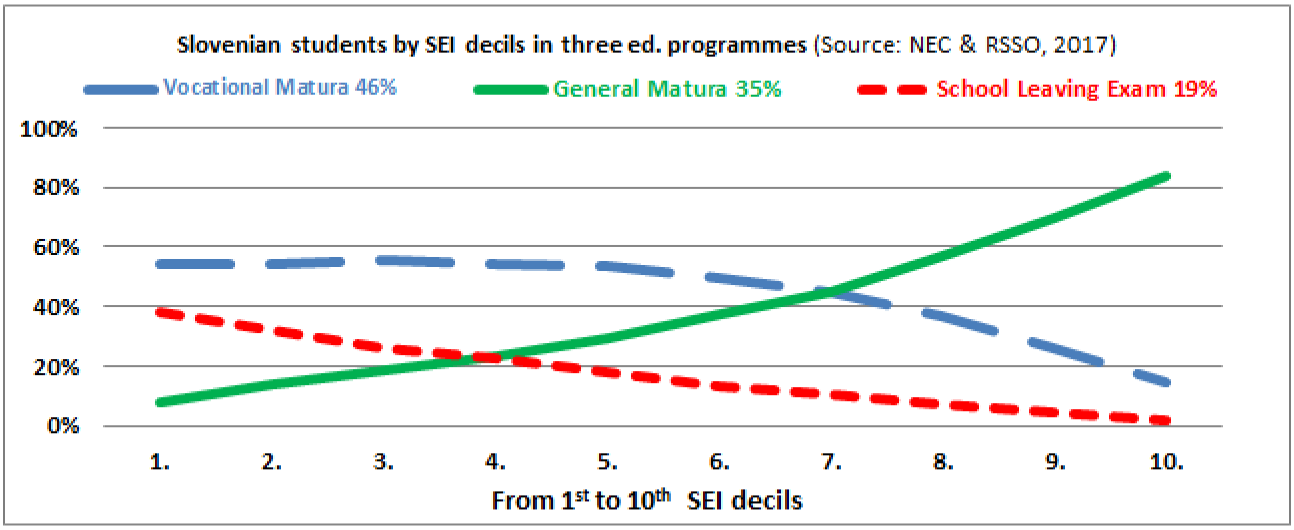
With the newly developed Social Economic Index (SEI) we can analyze associations of educational achievement of students in different upper secondary programmes in Slovenia with SES of students' families.
In the graph below, we present how the three upper secondary educational programmes with three different final exams (General Matura, Vocational Matura and School Leaving Exam) correspond with the socioeconomic index – in deciles.

Cankar, G., Bren, M. & Zupanc, D. (2017). Za večjo pravičnost šolskega sistema v Sloveniji (Analize povezav dosežkov učenk in učencev s socialnimi, kulturnimi, ekonomskimi in regionalnimi značilnostmi učenk in učencev, pridobljenih prek podatkov Statističnega urada RS). Ljubljana, Državni izpitni center.



A student from a family with SEI in the 10th decile in Slovenia will finish the upper secondary educational **Gimnazija program leading to General Matura** with 83.6% probability.
A student from a family with SEI in the 1st decile in Slovenia will finish the upper secondary educational **Gimnazija program** leading to General Matura with only 8% probability (Cankar, Bren, Zupanc, 2017).

A student from a family with SEI in the 10th decile in Slovenia will finish **upper secondary vocational education with a programme-ending exam** with only 2.0% probability.



The probability for students in short upper secondary vocational (2 years) and **upper secondary vocational education programmes** (3 years), where they need to pass a **programme-ending exam**, is reducing inversely proportional to the score of his/her family's SEI.

The probability for students in technical education programmes (**upper secondary technical & vocational technical education programmes**), where students need to pass the Vocational Matura, is constant up to 5th SEI decile; afterwards the probability is reducing when the family SEI goes from 6th to maximum 10th decile.

The higher is the SEI decile of the student's family, the higher the probability of the student finishing upper secondary education in Slovenia in the most demanding and prestige **Gimnazija programmes with General Matura Exam**, and the higher the probability of the student enrolling in tertiary education university programs (Cankar, Bren, Zupanc, 2017).

After an increase of enrolment in upper secondary education, there has been steady growth in the number of students, and a widening of expectations directed towards universities as well. However, **the growth in the number of students did not increase equity**, as it mostly favoured individuals from higher socio-economic groups or those whose parents with higher education.

A common assumption has been that a "free" system of higher education (one funded entirely by the state) is equitable by itself. In reality, this assumption did not materialize, since the main determining factor in participation is socioeconomic background. **Free access to higher education does not necessarily guarantee equity.** (COM; 2006: 7)

At this point, we can only speculate why free access to higher education does not guarantee equity. Probable causes may include:

- 'Hidden' costs involved. Families in an economically disadvantaged situation do not support their children in the transition to tertiary education, as this is still associated with high expenses they have to struggle with.
- Prolonged time of economic dependency. Either from the point of the students or from the point of their parents, higher education can be seen as an obstacle to their quicker (economic) independence. They would rather choose professions that require less education and lead to quicker employment, even if promised earnings are smaller.
- Social stereotyping. Up to this point in their educational path they have already internalized (or been indoctrinated with) the ideas of 'their place in society' and they do not identify with jobs that require higher education.
- Different orientation of values. They may simply view education as less important compared to other aspects of life. However, we can always question ourselves whether such values are indeed the cause for observed inequalities or the result.
- Influence of distorted media views. Media and public views of pessimism and negativity can create an impression that education is not important – what counts is instant (luck-based) success, and good jobs can only be had through connections with people in power, etc.

Most likely, there isn't just one isolated cause for observed inequalities (Voigt, 2017) but rather a myriad of influences. We must, however, ensure that the design of the educational system tries to minimalze those differences, not increase them.

Voigt, K. (2017). Distributive equality, relational equality and preferences about higher education. Theory and Research in Education, 15(2), 109‑128.
DOI: 10.1177/1477878517708116

X. Data-rich school environment

A data-rich school environment

It is difficult to have a highly functioning education system without a supportive institutional structure. Part of the reason for divergent opinions is simply a lack of sufficient experience, analysis, and evidence (Hanushek, Wössmann, 2011).

With regard to school systems, the OECD Centre for Educational Research and Innovation (CERI) promotes Educational longitudinal information systems:

- Use evaluation in a formative way for quick feedback to teachers, parents and students on students' performance;
- Put accountability and continuous improvement at the centre;
- Use it as knowledge management platforms for data driven decision-making; and
- enhance the precision of educational research and allow for more systematic evaluations.

http://www.oecd.org/edu/ceri/innovationstrategyforeducationandtraininglongitudinalinformationsystems.htm

Hanushek, E. A. in Wössmann, L. (2011). How much do educational outcomes matter in OECD countries? *Economic Policy, 26(67).* Pp. 427-491.

School Performance Feedback System (SPFS)

Within education systems, a culture of evaluation is needed. Effective long-term policies must be based on solid evidence.

For EU Member States to fully understand and monitor what is happening in their systems, they need channels for producing and accessing relevant research, a statistical infrastructure capable of collecting the necessary data, and mechanisms to assess progress as policies are implemented (COM, 2006).

Member States should develop a culture of evaluation. They should develop policies for the entire lifelong learning continuum which take full account of effectiveness and equity in combination and in the long term, and which complement policies in related fields.

As data based decision-making is receiving increased attention in education, more and more **school performance feedback systems (SPFS)** are being developed and used worldwide. These systems provide schools with data on their functioning. (Verhaeghe, Schildkamp, Luyten, Valcke, 2015).

Many schools in countries all over the world use school performance feedback systems (SPFS) to gather these data, which "are information systems external to schools that provide them with confidential information on their performance and functioning as a basis for school self-evaluation" (Visscher & Coe, 2002).

The Slovenian National Examinations Centre (NEC) also developed a version of SPFS for purposes of ensuring effectiveness and attaining improvement in classrooms and schools in upper secondary education in Slovenia: Assessment of/for Learning Analytic Tool (ALAT) (Zupanc, Urank, Bren, 2009).

Software enables interactive analysis of data for each school with national benchmarks, calculated using the same selection criteria (Brejc, Sardoč & Zupanc, 2011). In 2014, the software was upgraded with new features: per item analysis for criterion-referenced interpretation and value added analysis.

Teachers, head teachers and other educators in Slovenia are able to access the examination database, analyze assessment results and teachers' grades, interpret the achievement of their students, and analyze the efficiency of teaching and learning in classrooms and schools in pre-university upper secondary schools with the General and Vocational Matura as the exit examinations.

Schools can use the Slovene 'educational longitudinal information system' for self-evaluation and data-driven decisions on improvement in teaching and students' learning, and for higher level of equity in schools.

Verhaeghe, G., Schildkamp, K., Luyten, H., & Valcke, M. (2015). Diversity in school performance feedback systems. *School effectiveness and school improvement, 26 (4)*. 612 - 638. ISSN 0924-3453
Visscher, A. J., & Coe, R. (Eds.). (2002). School improvement through performance feedback. Lisse: Swets & Zeitlinger.
Zupanc, D., Urank, M., & Bren, M. (2009). Variability analysis for effectiveness and improvement in classrooms and schools in upper secondary education in Slovenia: Assessment of/for Learning Analytic Tool. *School Effectiveness and School Improvement, 20*, 89–122.
Brejc, M., M. Sardoč & D. Zupanc (2011). *Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Slovenia.* National School for Leadership in Education in cooperation with the National Examinations Centre, http://www.oecd.org/education/school/48853911.pdf: 54–55.

Joke: *"I'm making a decision! Stop confusing me with facts!"*

1. Slovene ALA Tool as School Performance Feedback System (SPFS)

In 2007, National Examinations Centre (NEC) developed the Slovene **School Performance Feedback System (SPFS)** – the Assessment of/for Learning Analytic (ALA) Tool (Zupanc, Urank, Bren, 2009).

As a school performance feedback system, ALA Tool allows all upper secondary schools interactive analysis of their results in compulsory external examinations taken at the end of schooling. Stakeholders that can use the software directly are mostly head teachers and teachers who, through recent improvements, can compare results of their students with national benchmarks to draw conclusions for teaching.

In 2012, ALA Tool was upgraded with Value Added measures and options for criterion-referenced interpretations of external assessment outcomes. The value-added approach enables schools to see the progress of their students from the end of basic school to the end of upper secondary school.

ALA Tool gives new applicability to Matura results for ongoing analyses of student assessment data, monitoring and evaluation of the school system at national, school and classroom level for several consecutive years (2002–2017). It includes the achievement of the sixteen-year cohorts, i.e. 296,000 secondary school students. This is almost 15% of the entire Slovene population.

The Tool provides continuous follow-up of schools and classes (teachers) in schools regarding the effectiveness and school improvement on a more empirical, data-driven basis, i.e. useful Management Tool in Education. ALA Tool is used for decisions on improvement in teaching and students' learning, and for a higher level of equity within and between schools.

The system provides integrated tools to enable easy data comparisons. The system could also be applied to analyze the variability on different levels of the school system in different subjects.

Under the project Innovation Strategy for Education and Training, the Centre for Educational Research and Innovation (CERI) is greatly interested in the existence of longitudinal data systems – also in ALA Tool in Slovenia.

With this online Tool, head teachers and teachers in all secondary schools in Slovenia are able to access the examination database and analyze the achievements of their students. They may compare them to peer groups outside of school and to classes within a school, and recognize trends.

The main goals of the ALA Tool are:
- support management at classroom, school and national level for self-evaluation and decision making;
- compare a class or a school to other schools sharing common characteristics;
- compare students' achievements between classes (teachers) within a school;
- analyse possible drop-out rates;

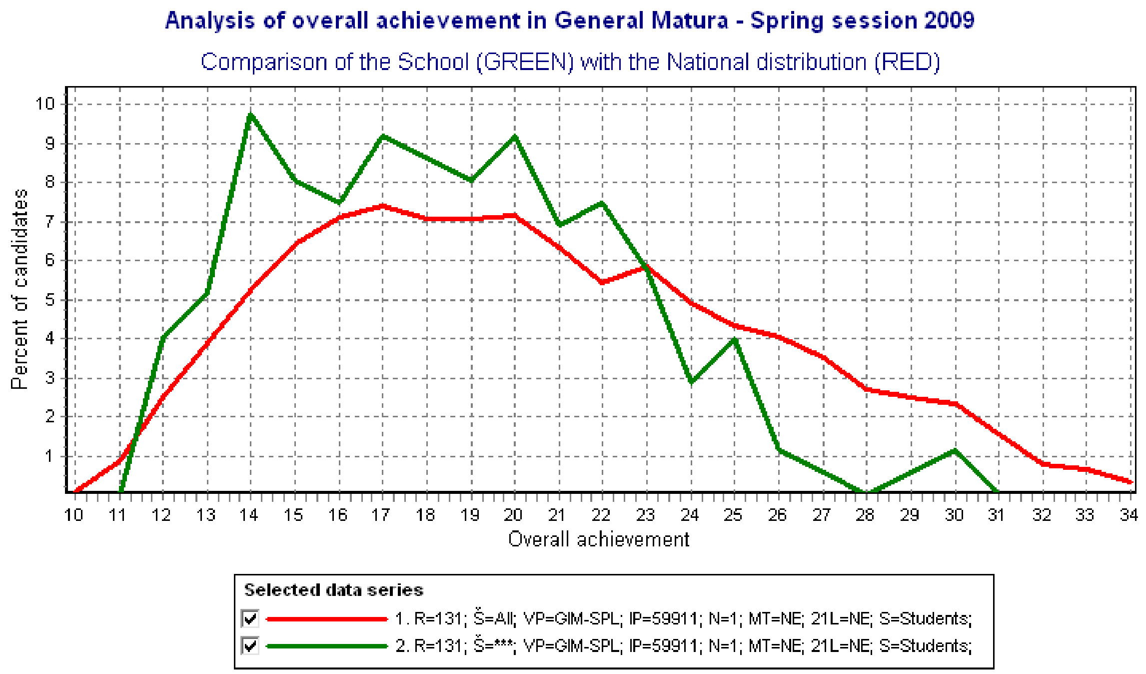
•recognise longitudinal trends;
•analyse teachers‘ internal grading and external marking;
•analyse students' selection of different optional subjects in classrooms and schools and selection of different tiers of compulsory subject exams;
•compare outcomes in one subject to another and to overall achievement;
•analyse students' outcomes in written and oral assessment, coursework;
•analyse educational progress of student with value-added measures;
•analyse achieved objectives and performance standards with items analysis.

The tool for the analysis of the demonstrated skills at the end of upper secondary education allows schools to separate and filter the data by various criteria, for example sex, class, course of study etc. The flexibility of the software provides schools with detailed insight into their students' performance and a comparison with the corresponding groups of students and schools at national level.

In particular, using multi-annual databases of students' achievements help schools, as well as teachers, to analyse and interpret students' achievements; compare and analyse differences in achievements between classrooms and teachers in their school; compare the results of different forms of assessment: teacher's grade, external grade, written and oral grades, coursework grade, etc.

ALA Tool presents Slovenian Educational Information Systems for innovation and improvement in education.

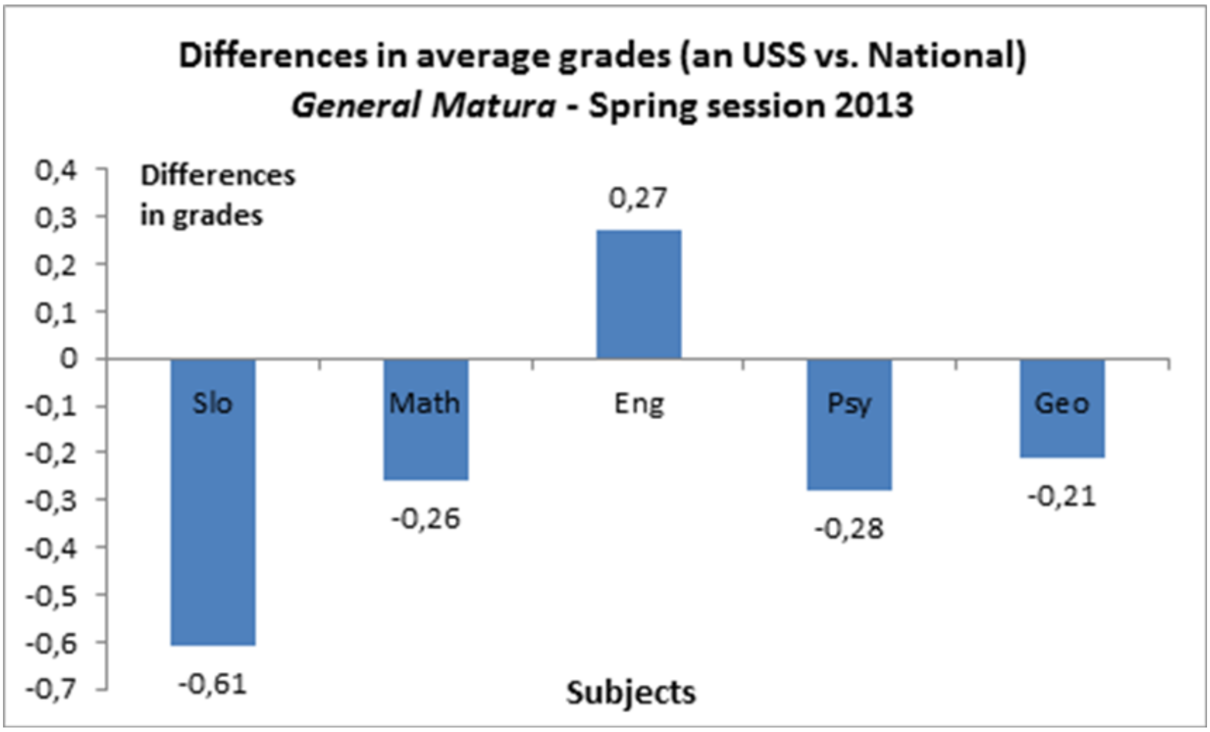
For instance: ALA Tool allows each upper secondary school (USS) with General Matura or Vocational Matura to compare their overall achievement in the Matura session with the overall achievement of all schools in Slovenia with the same upper secondary program type. In the graph bellow, there is a distribution of overall achievement (a USS vs. National distribution) for spring session of General Matura 2009.



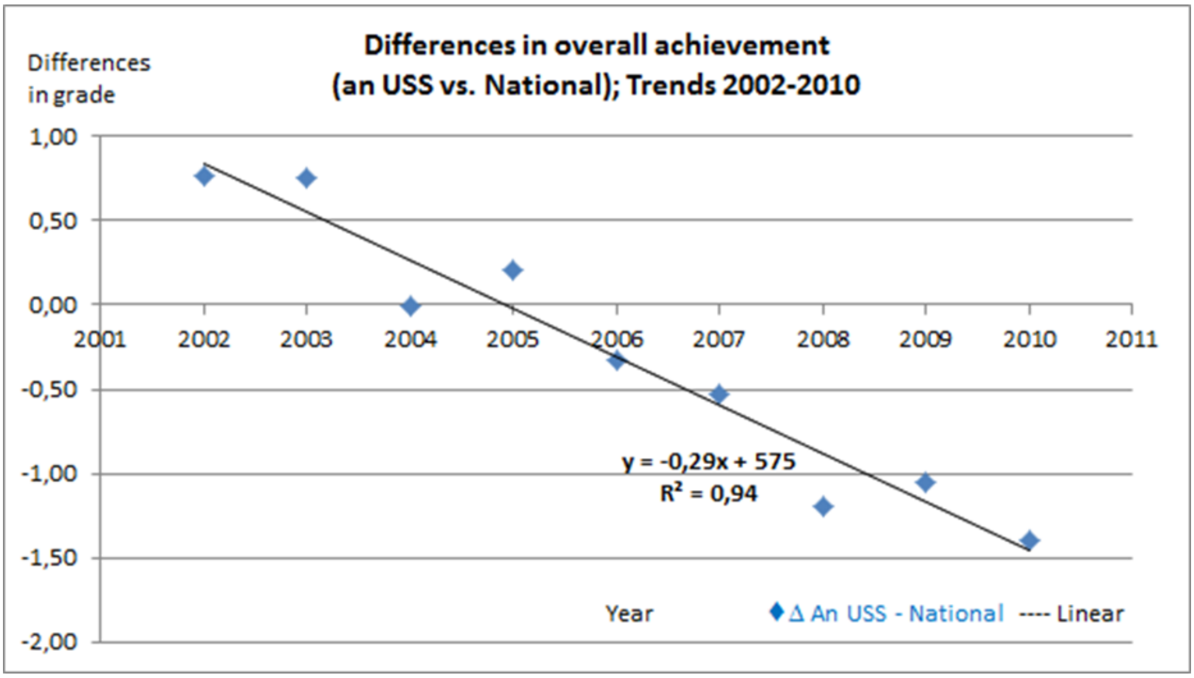
Basic statistics
1: No. cand.: 5049; No. schools: 55; Average: 20,41; STD: 5,14
2: No. cand.: 174; No. schools: 1; Average: 18,48; STD: 3,93

So-called "public opinion" or rumors say that one school is above-average and another school is below the national average... There are differences in student achievements between subjects within a school, and even between teachers in one subject within a school. It often happens that achievements in some subjects in a school are above average, whereas achievements in other subjects are below the national average.

The graph below shows the difference in average grades in one upper secondary school (a USS vs. National) in the spring session of General Matura 2013

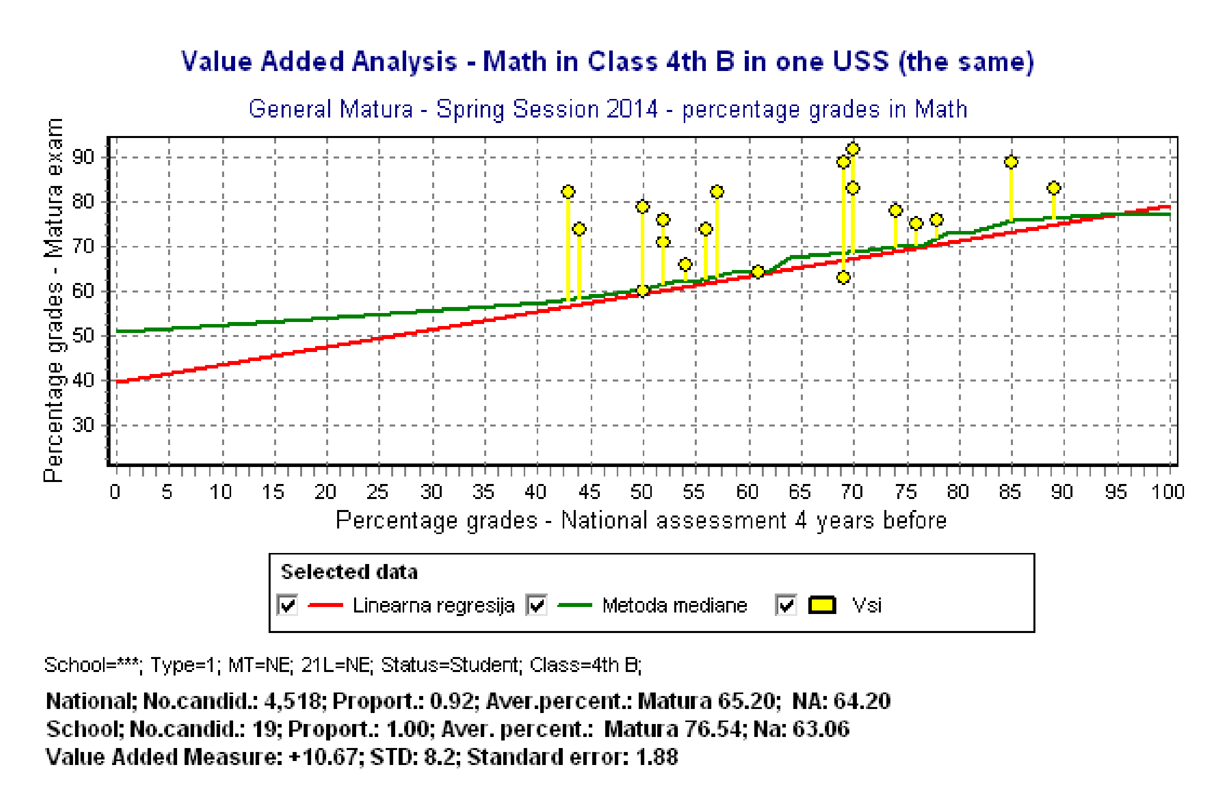


Schools and teachers use multi-annual databases of students' achievements to analyse longitudinal trends in each Matura subject and in overall achievements. The graph below shows the difference in overall achievement between a school and the national average – trends. It is an extreme case, but it is real.



The NEC has results of the National Assessment (NA) at the end of compulsory education – in the last year (9th grade) of basic school. There are data for achievement of Mother Tongue, Math and 1 optional subject for each student in the generation in Slovenia. Value-added approach enables schools to analyse the progress of their students from the end of basic school to the end of upper secondary school.

The graph below shows the value-added analysis for a Math teacher in Class 4th B in an upper secondary school in spring session of General Matura 2014. The subject is an excellent teacher; almost all his/her students progress better than their peers with similar NA results (four years before Matura Math exam).



XI. Caution due to misuse of data

1. Example 1: PISA and equity indicator – lesson from Slovenia

The Programme for International Student Assessment (PISA) is an internationally renowned large-scale assessment that measures achievement of 15 year olds who are still in formal education in three-year cycles in the field of reading literacy, mathematical literacy and science literacy (OECD, 2007). It provides a wealth of high-quality data for researchers in education and is a driving force for many educational reforms worldwide. It stands to reason that the results of PISA are highly regarded, and many questions were raised when Slovenia got very low results on the equity indicator in PISA 2006. As shown in Figure 1, Slovenia ranked third highest on the between-school variance indicator – an indicator of great differences between schools. Over 60% of all differences in science achievement of students was attributable to schools! In other words, in Slovenia students were attending schools that (over time) created remarkable differences between students, which leads to the conclusion that some students were in 'good' schools and some students in 'bad' ones and the educational system was promoting this.

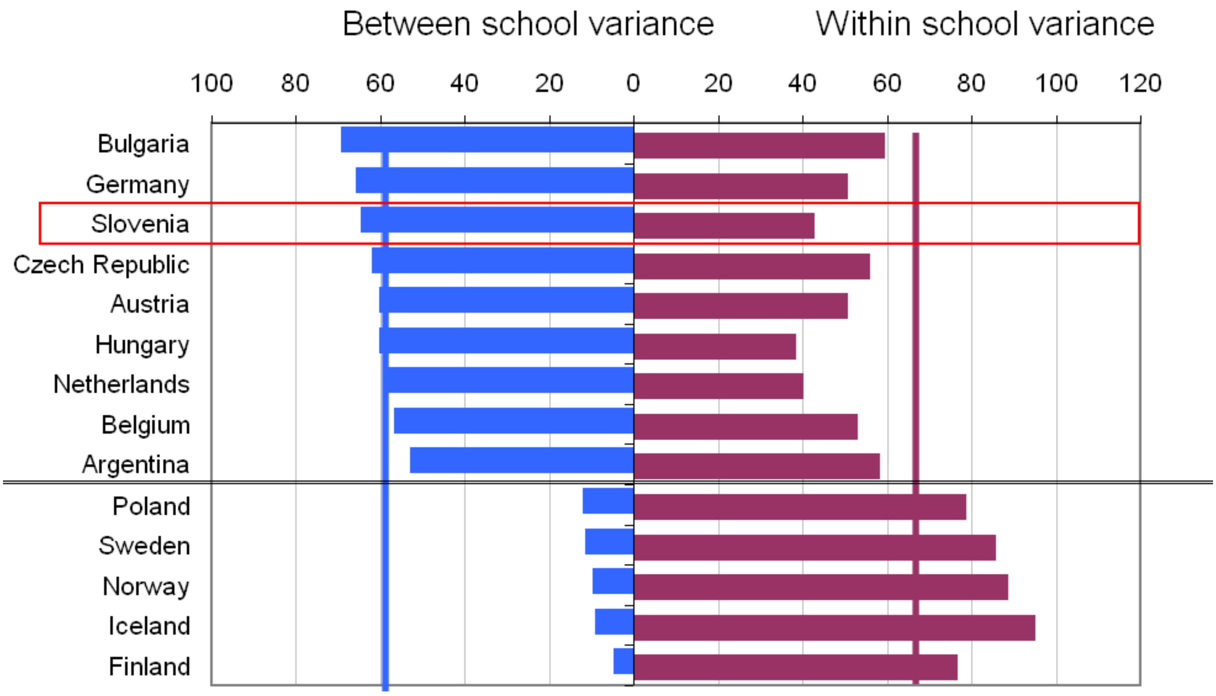


Figure 1: Excerpt from results of PISA 2006 (OECD, 2007; p171).

The results were inconsistent even at a quick glance. Researchers should be neutral and accept the results as they are, but what if the results do not fit common sense? In Slovenia, students in compulsory schooling (first nine years) go to a nearby school – there are local catchment areas defined. As Slovenia is not a highly segregated country, these areas are not so different that they could explain the third worst result between all the countries in PISA.

After careful analysis, we established that there was a problem with the indicator of equity. In Slovenia, 15-year-old students are in the vast majority in their first year of upper secondary schools. They are in schools that are not accountable for the knowledge PISA is measuring, and since they just decided where their educational career path will lead them, schools very much differ in their respective composition of students.

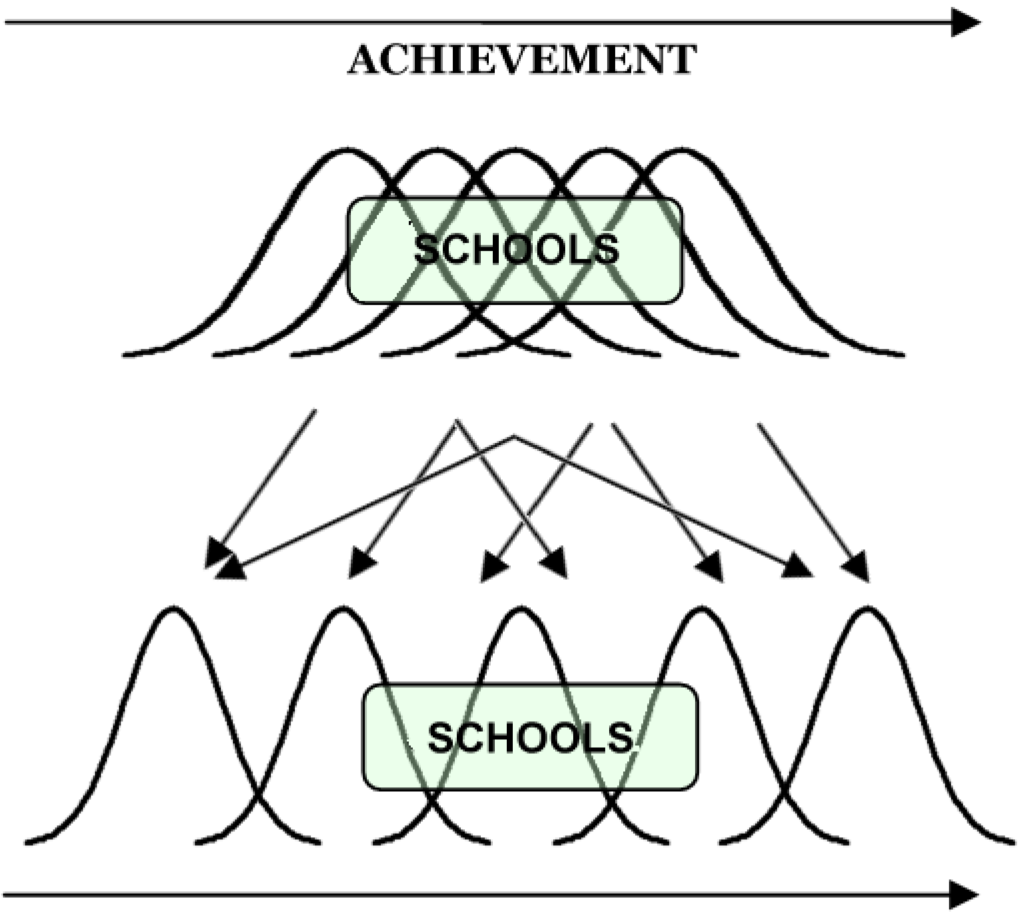


Figure 2: Students left primary schools and went to upper secondary schools that were much more different regarding the composition of students.

It was actually a case of mixed identity (of schools). Results were calculated according to upper secondary schools (students attended at the time of PISA testing), but findings were interpreted as results of previous nine years of schooling!

As we have at our disposal national assessment at the end of the nine-year school cycle (programmes), we were able to compare PISA results with national data. As shown in Table 1, proportions of between-school variance were much lower than PISA reported across different schools' subjects and years, and although they were not perfect – equity could still be improved – they could be interpreted as more valid than PISA results (Gaber, Cankar, Marjanovič Umek & Tašner, 2012).

Table 1: Proportions of between-schools variance for national assessment data

Subject	Year	No. of students	No. of schools	Between schools	Within schools
Slovene	2006	20,689	477	15.90	84.10
Slovene	2007	19,056	478	15.54	84.46
Mathematics	2006	20,832	479	13.21	86.79
Mathematics	2007	19,311	480	13.11	86.89
Biology	2006	20,833	480	10.69	89.31
Physics	2007	4,548	114	13.65	86.35
Chemistry	2007	4,937	121	20.12	79.88

GABER, Slavko, CANKAR, Gašper, MARJANOVIČ UMEK, Ljubica, TAŠNER, Veronika. The danger of inadequate conceptualisation in PISA for education policy. *Compare*, ISSN 0305-7925, 2012, vol. 42, iss. 4, str. 647-663. http://www.tandfonline.com/doi/abs/10.1080/03057925.2012.658275, doi: 10.1080/03057925.2012.658275.

OECD. (2007). *PISA 2006 Science Competencies for Tomorrow's World, Volume 1: Analysis*. OECD.

2. Example 2: Differences between students in general and vocational Matura

In Slovenia, upper secondary schools mainly fall into two major tracks – schools that end in General Matura and lead to universities, and schools that end in Vocational Matura and provide vocational qualifications. However, students in programmes ending in Vocational Matura can apply for tertiary education as well, and there's no established measure how different examinations in General and Vocational Matura actually are in terms of difficulty.

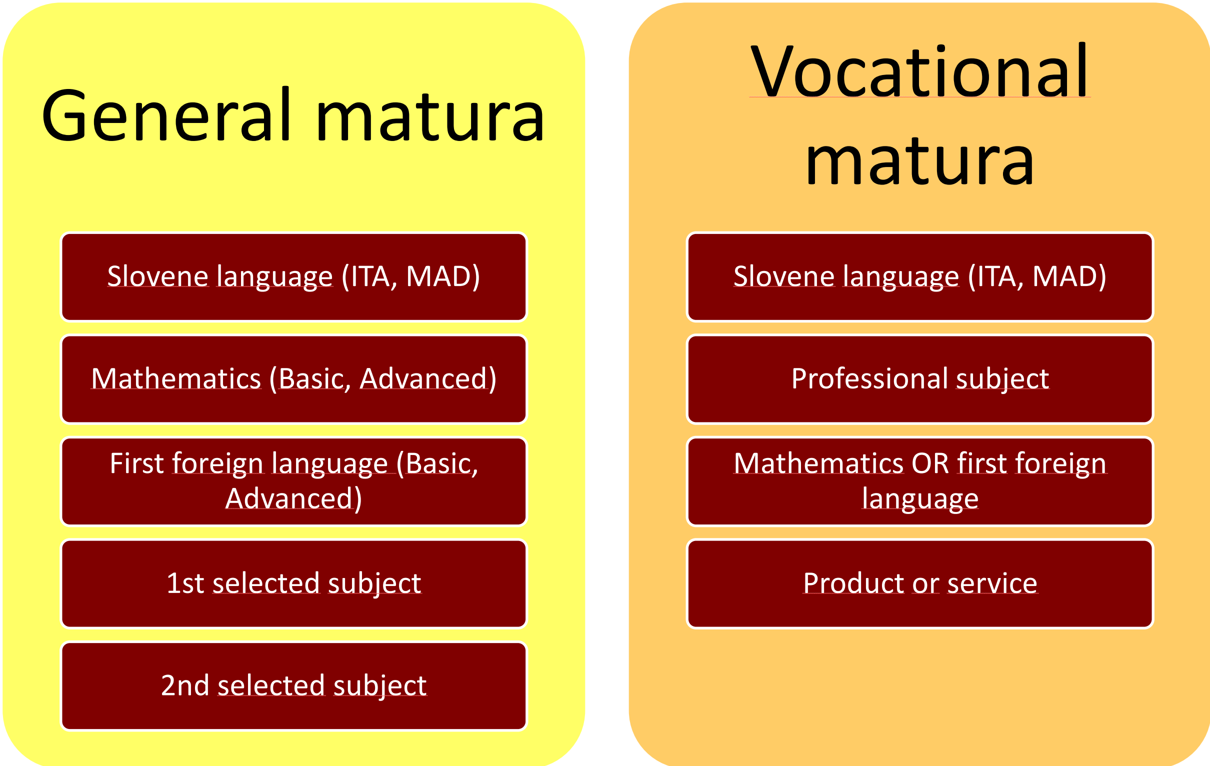


Figure 1: Comparison of General and Vocational Matura in Slovenia

General and Vocational Matura share many similarities, but there are also major differences. General Matura has five examinations, Vocational only four, mathematics is obligatory in General Matura, but only optional in Vocational, etc. There are also differences in implementation and scoring.

However, results from both are applied during admission to university programmes, and sometimes grades from individual subjects count as well. A good example is mathematics, a subject appearing both in General and Vocational Matura. Since there are no detailed studies about equivalence of grades from General and Vocational Matura, Slovenian admissions offices use grades from both as equivalent. Naturally, the question arises whether this is fair.

Here we provide partial evidence that this cannot be the best solution available. We cannot put results of students from General and Vocational Matura on a common denominator scale – students had different examinations and got points on different scales. However, we do have results from all students before they entered upper secondary education. Results of the national mathematics examination students took just before they entered their General or Vocational education track can be compared with the grade they received in mathematics at each Matura.

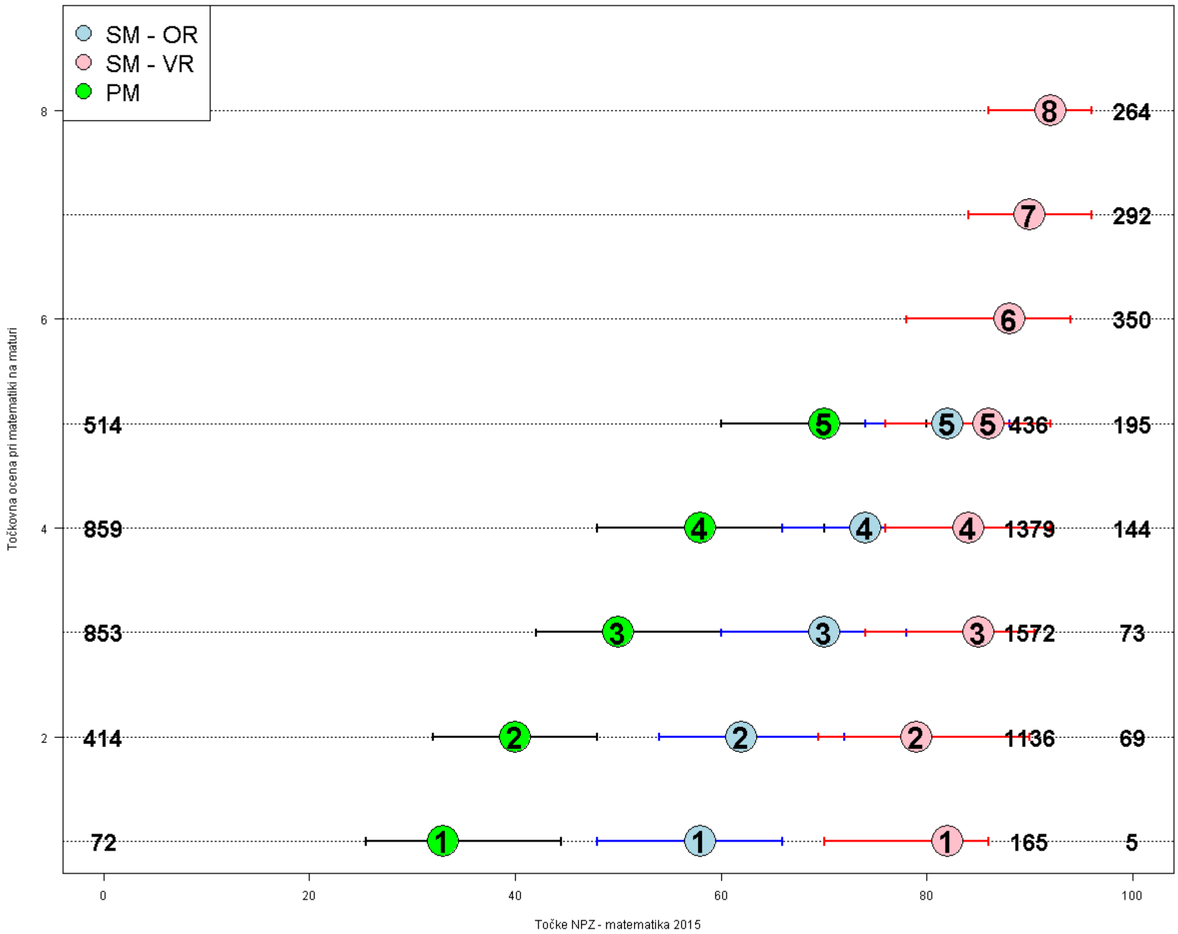


Figure 3: Grades from Mathematics at General Matura (blue and pink for basic and advanced level, respectively) and Vocational Matura (green), compared to results from national assessment in Mathematics four years prior. Circles represent median scores while bars represent range of students' achievement.

Figure 3 shows the skills of students four years prior, and the grades they received at Matura examinations. For example, students who received a grade of 5 in Mathematics at Vocational Matura are much more similar in their composition to the students who got 4s or even 3s in Mathematics at General Matura. From this data, we cannot create a table of equivalent grades, since our conclusion is based on data gathered four years prior to Matura results. Results of a more complex and detailed study that would gather additional data could provide better evidence. However, given the fact that the educational track ending in General Matura has a more demanding curriculum, those differences had most likely only grown bigger over time.

3. Example 3: Case of PISA results and Simpson's paradox

This is a simple case of misinterpretation in statistics. It is a simple lesson that sometimes nothing is wrong with the data, but we must still be careful when we interpret it. We will again use the results of PISA (2006 cycle (OECD, 2007)).

Table 1: Results of science achievement on PISA 2009 by educational track and gender

SLOVENIA

Ed. tracks	Female	Male	Difference	Total
2 year voc.	354	365	-11	363
3 year voc.	406	429	-23	421
4 year tech.	483	504	-21	495
4 year tech. Gymn.	528	560	-32	544
4 year gen. Gymn.	590	614	-25	599
Total	526	519	7	523

The first shows different educational tracks where we find students aged 15 in Slovenia. Educational tracks are listed in order according to the difficulty of their curriculum, starting with 2-year vocational education courses that train for simple vocations, and ending with the General Gymnasium track that lasts 4 years and is considered most demanding track in Slovenia that leads to all university courses directly.

Table 1 shows interesting results. In each educational track, boys outperform girls in their science achievements by 10, 20 or even more than 30 points. These differences are statistically and practically significant. Naturally, you would expect that boys generally outperform girls in science achievements. However, the results negate this conclusion – girls outperform boys by 7 points! How is that even possible?

There is nothing wrong with the student sample, calculations made or statistics used. The results are the consequence of differences in proportions of boys and girls in those educational tracks.

Figure 4 shows proportions of students by gender in each of the tracks. Tracks are listed in order from most demanding to least demanding. Now, every student (a boy or a girl) wants to go to the most demanding track s/he thinks s/he can manage. While their abilities for science are not that different between boys and girls, their preference for educational tracks are. In the most demanding, i.e. demanding 4-year general Gymnasium track, there are fewer boys than girls. Since this group of boys presents a "more talented selection" compared to their female peers, their achievement is higher than the achievement of girls in the same track. In the next educational track (4-year technical Gymnasiums), the remaining boys still outperform the (remaining) girls for similar reasons. This continues across each educational track until the last one, where boys are in the majority and they outperform the smaller population of girls (as girls with higher science achievement already went to other tracks).

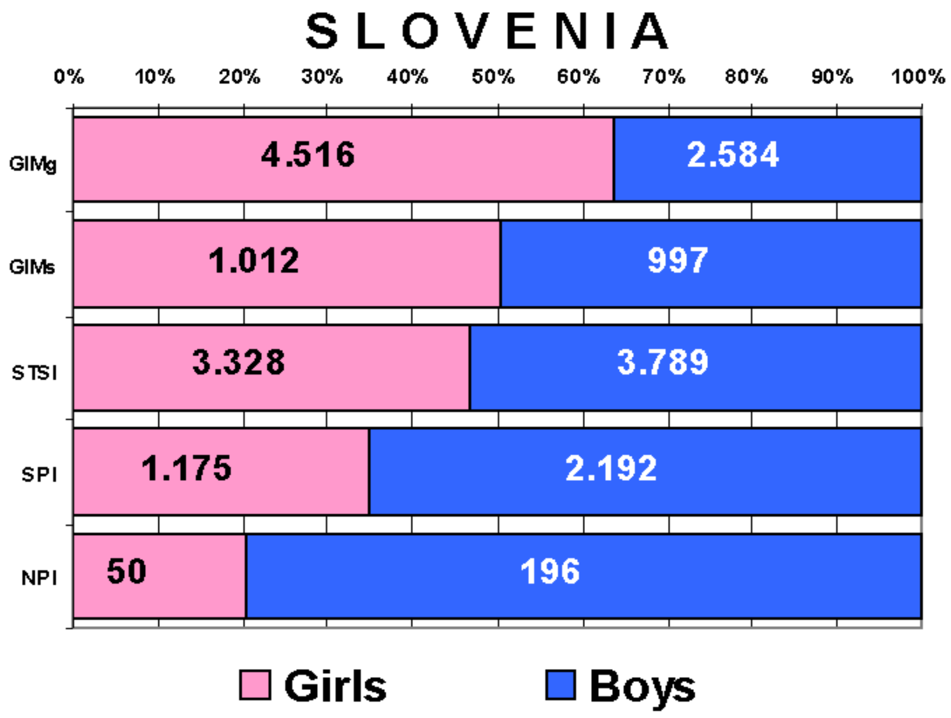


Figure 4: Proportions of boys and girls by educational tracks. Educational tracks are listed by difficulty of curriculum (GIMg – general Gymnasium , GIMs – technical Gymasium, STSI – 4 year vocational track, SPI – 3 year vocational track, NPI – 2 year vocational track).

In literature, this is known as Simson's paradox (Blyth, 1972) and it teaches us to watch out for underlying proportions of subpopulations not only their results, as they can be misleading.

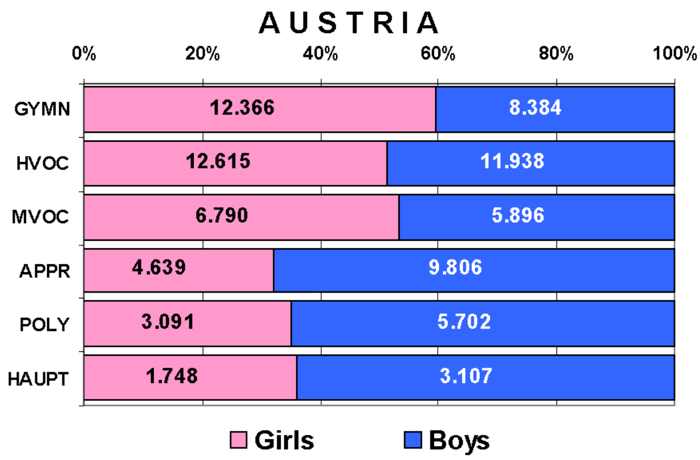
We can observe identical phenomena in many other countries as well – here we provide tables and figures for some other countries. In case of Croatia, there is a sign of the difference between boys and girls having changed (similar as in Slovenia), whereas with other countries, differences are just diminished.

This has practical implications as well – since science skills is not the only criterion influencing choice of educational track, many students that are skilled in science decide to remain in less demanding educational tracks.

Data also suggest that systems create a self-fulfilling prophecy – albeit boys and girls start with science achievements that are more similar, we guide selection to upper secondary schools in such a way that data confirm the stereotype image that boys are better at science than girls.

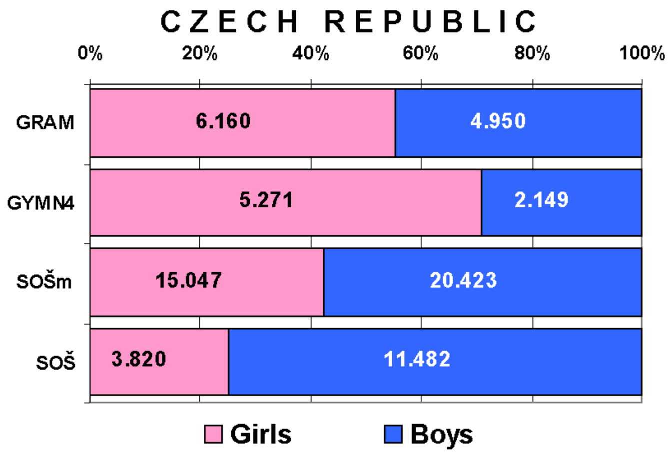
AUSTRIA

Stratum	Female	Male	Difference	Total
GYMN	574	584	**-9**	578
HVOC	544	573	**-30**	558
MVOC	442	500	**-58**	469
APPR	435	472	**-37**	460
POLY	418	447	**-30**	437
HAUPT	386	414	**-28**	404
Total	508	516	**-9**	512



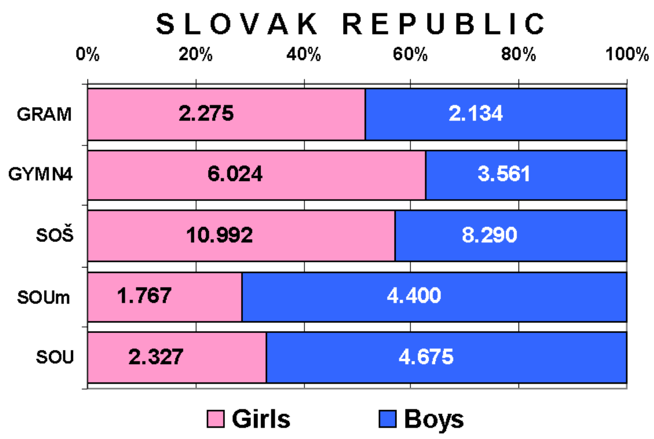
CZECH REPUBLIC

Stratum	Female	Male	Difference	Total
GRAM	617	642	**-25**	628
GYMN4	609	625	**-16**	613
SOŠm	525	556	**-31**	542
ZAKL	481	494	**-13**	489
SOŠ	413	453	**-40**	443
Total	516	520	**-3**	518



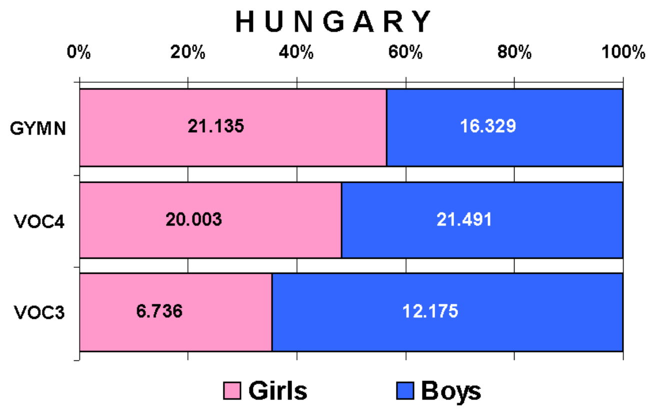
SLOVAKIA

Stratum	Female	Male	Difference	Total
GRAM	584	601	**-17**	592
GYMN4	554	580	**-26**	564
SOŠ	488	521	**-33**	502
SOUm	449	483	**-34**	473
ZAKL	457	469	**-13**	464
SOU	407	420	**-13**	416
Total	487	494	**-7**	490



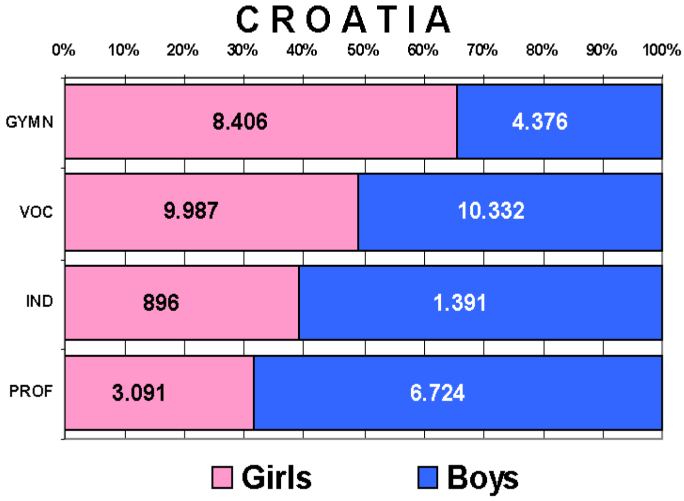
HUNGARY

Stratum	Female	Male	Difference	Total
GYMN	550	576	**-27**	561
VOC4	492	522	**-30**	507
VOC3	419	437	**-19**	431
Total	507	519	**-12**	513



CROATIA

Stratum	Female	Male	Difference	Total
GYMN	554	584	**-30**	565
VOC	479	507	**-28**	493
IND	410	429	**-19**	421
PROF	408	425	**-18**	420
Total	495	493	**2**	494



Tables show PISA 2006 science achievements and proportions of boys and girls by educational tracks in Austria, the Czech Republic, Slovakia, Hungary and Croatia.

Blyth, C. R. (1972). On Simpson's Paradox and the Sure-Thing Principle. *Journal of the American Statistical Association*, *67*(338), 364-366.
OECD. (2007). *PISA 2006 Science Competencies for Tomorrow's World, Volume 1: Analysis*. Paris: OECD.

XII. Pre-university TIMSS Advanced study

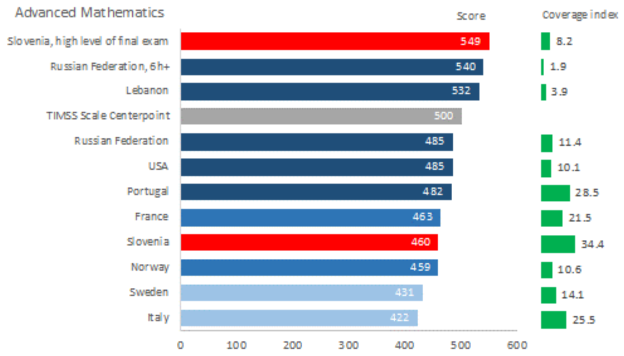
Pre-university TIMSS Advanced study 1995, 2008 and 2015

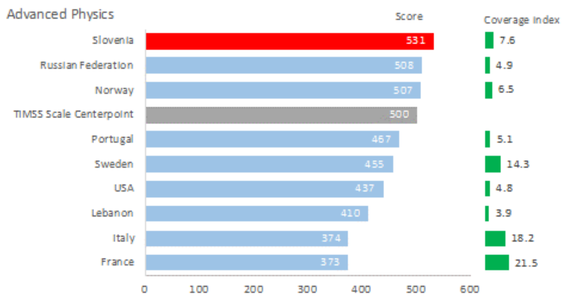
TIMSS Advanced is the only international assessment that provides essential information about students' achievement in advanced mathematics and physics.
It assesses students in their final year of secondary school (often 12th grade) who are engaged in advanced mathematics and physics studies that prepare them to enter STEM tertiary programs (science, technology, engineering, and mathematics) in higher education.
TIMSS Advanced was first conducted in 1995 as part of TIMSS and then separately again in 2008 and in 2015.
9 countries participate in TIMSS Advanced 2015: France, Italy, Lebanon, Norway, Portugal, the Russian Federation, Slovenia, Sweden, and the United States; 7 out of 9 participated also in 1995 and 6 out of 9 participated in 2008.

IEA (2016). TIMSS ADVANCED 2015 INTERNATIONAL RESULTS IN ADVANCED MATHEMATICS AND PHYSICS; http://timss2015.org/advanced/

Graphs below show the Average Advanced Mathematics Achievement (Score) and Advanced Mathematics Coverage Index for 9 countries that participated in TIMSS Advanced 2015

The advanced mathematics or physics coverage index is the percentage of the corresponding age cohort covered by students in their final year of high school who are taking or have taken advanced mathematics or physics courses. The corresponding age cohort is determined for education systems individually.





Barbara Japelj Pavešić, TIMSS Advanced National Research Coordinator from Educational Research Institute in Ljubljana commented the Slovene results (Japelj, 2016; Youtube):

Regarding advanced mathematics, in our country, mathematics is a compulsory subject for all future university students, independent of their future areas of study at university.

Mathematics exam at the end of the general secondary school is compulsory for all. It can be taken at basic or high level.

Score is used as criterion for admission to the university studies with limited entrance.

Students who want to enter these studies, such as medicine or biomedical studies, are highly motivated for high math score. In TIMSS Advanced, they reached the highest mathematics score from all populations, as seen on the graph.

Most STEM studies with limited entrance also required candidates to pass the final exam in at least one science subject. Therefore, students who want to enter limited STEM studies are also highly motivated for high physics score.

Our specialists, dr. Zupanc, Director of the Examination Center, and dr. Planinšič, from the Faculty or Mathematics and Physics, who train future physics teachers, agree that the important reason for high math knowledge in our country are well educated one-subject teachers of advanced mathematics and physics.

In Slovenia, a teacher of advanced mathematics or physics can become a person who finishes 4 to 5-year university study of mathematics or physics, with additional pedagogical courses.

There was also implemented and then revised curriculum for all grades 1 to 13, between 2008 and 2011, with improved definitions of topics and standards of knowledge that could contribute to good results.

In compulsory schools, we observe the highest increase of results in grade 4. Teachers at that level are class teachers. In last years, the university study for class teachers became more popular, even it has limitation of entrance. As a consequence, population of future teachers improved and therefore also population of teachers.

The important reason for improvement are also TIMSS results. They were always well accepted and used in all levels of the system. Teachers use TIMSS items to encourage reasoning.

Video can be found at https://www.youtube.com/watch?v=J4BHKTWDwgU; go to 43:35 and follow until 45:50

Admission to higher education rules and good results in TIMSS Advanced in Slovenia (Zupanc, 2016; Youtube)

In Slovenia, the math and science curriculum and universal standards are well aligned[1] with the national-assessment[2] and Curriculum-Based External Examination System (CBEES).[3]

Passing the Matura exam with compulsory math proves general ability for any academic course.

The math score is even more important for students applying to the most prestigious faculties that have enrolment thresholds.

According to the Rules on admission to higher education, these students should take math at the most advanced level, and many of them should prove excellence by taking physics or chemistry or biology or two of them[4] as additional Matura subjects.

With well-educated math and science teachers and hard work, a large percentage of students can achieve high standards.

For effectiveness and improvement purposes, teachers and school leaders in Slovenia have access to *School Performance Feedback System* (*SPFS*)[5] – *Assessment for/of Learning Analytic Tool* (*ALA Tool*)[6] – for interactive confidential analysis of long-term students' achievement.

1 Linn, R.L. (2005), "Issues in the Design of Accountability Systems", University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles. Looney, J. (2009). *Assessment and Innovation in Education*. OECD Education Working Paper No. 24. Pridobljeno 16. 7. 2009na http://ideas.repec.org/p/oec/eduaab/24-en.html

2 BAJC, J. & PAVLIN, J. (2010). National assessment of knowledge of physics and curriculum development in Slovenia. In: DOLINŠEK, Slavko (ed.). *Socio-cultural and human values in science and technology education : proceedings*, XIV. IOSTE Symposium, International Organization for Science and Technology Education, June 13.-18. 2010, Bled, Slovenia. Ljubljana: Institute for Innovation and Development of University, 2010, pp. 1-7, http://www.ioste.org/pdf/proceed14.pdf.

3 Bishop, J. (1998), "Do Curriculum-based External Exit Exam Systems Enhance Student Achievement?", CPRE Research Report Series RR-40, Consortium for Policy Research in Education, University of Pennsylvania, Philadelphia.
&
Bishop, J. (2001), "How External Exit Exams Spur Achievement", in F. Mane and M. Bishop (eds.) *Educational Leadership,* Association for Supervision and Curriculum Development, Baltimore, Maryland.

4 Zupanc, D. in Bren, M. (2009). Optional subjects in the matura exam: are candidates who select science subjects higher achievers? V: 35th Annual Conference Assessment for a creative world, Brisbane, 13.–18. September 2009. *35th Annual Conference Assessment for a creative world: Brisbane, Australia, 13.–18. September 2009*. Brisbane, 2009, p. 1–11, http://www.iaea.info/documents/paper_4d73f71.PDF.

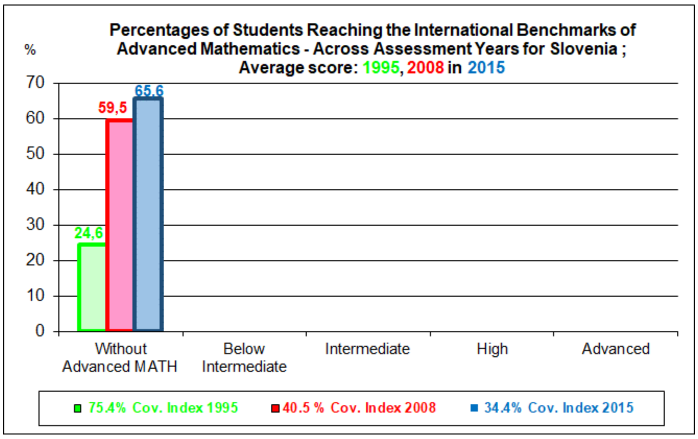
5 Visscher, A.J. & Coe, R. (2003). School Performance Feedback Systems: Conceptualisation, Analysis, and Reflection. *An International Journal of Research, Policy and Practice, 14(3),* 321-349.

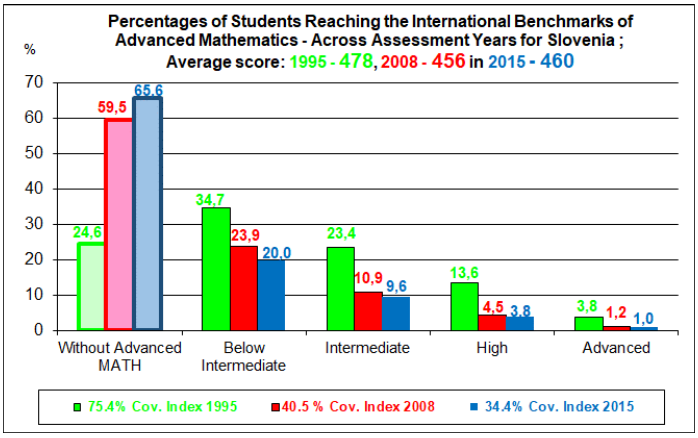
6 Zupanc, D., Urank, M. in Bren, M. (2009). Variability analysis for effectiveness and improvement in classrooms and schools in upper secondary education in Slovenia: assessment of/for learning analytic tool. *School Effectiveness and School Improvement, 20(1), 89-122.*
&
ZUPANC, D. & CANKAR, G. (2015). Upgrade of slovenian school performance feedback system (SPFS) with individual's contextual data - feasibility study. V: ECER 2015 : education and transition - contributions from educational research : programme by network. Budapest: European Educational Research Association, 2015, http://www.eera-ecer.de/ecer-programmes/conference/20/contribution/34128/.

The graph below shows the differences in Advanced Mathematics achievement across assessment years for Slovenia. The graph also shows the percentages of students reaching the international benchmarks of Advanced Mathematics achievement across assessment years for Slovenia (IEA, 2016; Exhibit M2.3.)

TIMSS Advanced describes achievements at three International Benchmarks along the scale: **Advanced**, **High** and **Intermediate**. The difference of the coverage index to 100 % presents the proportion of those **Without advanced mathematics**.

Not only do the graphs show apparent decline in achievements over the years, they also show that certain proportions of Advanced Mathematics were reached by a much smaller proportion of the generation in the past year.





TIMSS Advanced describes achievement at three International Benchmarks along the scale: **Advanced**, **High**, and **Intermediate**. There are descriptions of performance at the International benchmarks of Advanced Mathematics Achievement (IEA, 2016; Exhibit M2.2: for 2015)

Intermediate Benchmark
Students demonstrate basic knowledge of concepts and procedures in algebra, calculus, and geometry to solve routine problems.
Students can apply and transform a formula to solve a word problem. They can determine a term in a geometric sequence and analyse a proposed solution of a simple logarithmic equation. They can recognize a graph of the absolute value of a function, and identify and evaluate composite functions.
Students can find the derivative of exponential, trigonometric, and simple rational functions. They can find limits of rational and exponential functions. They can make connections between the sign of the derivative and the graph of a function.
Students can use knowledge of basic properties of geometric figures and the Pythagorean Theorem to solve problems. They can add and subtract vectors in coordinate form.

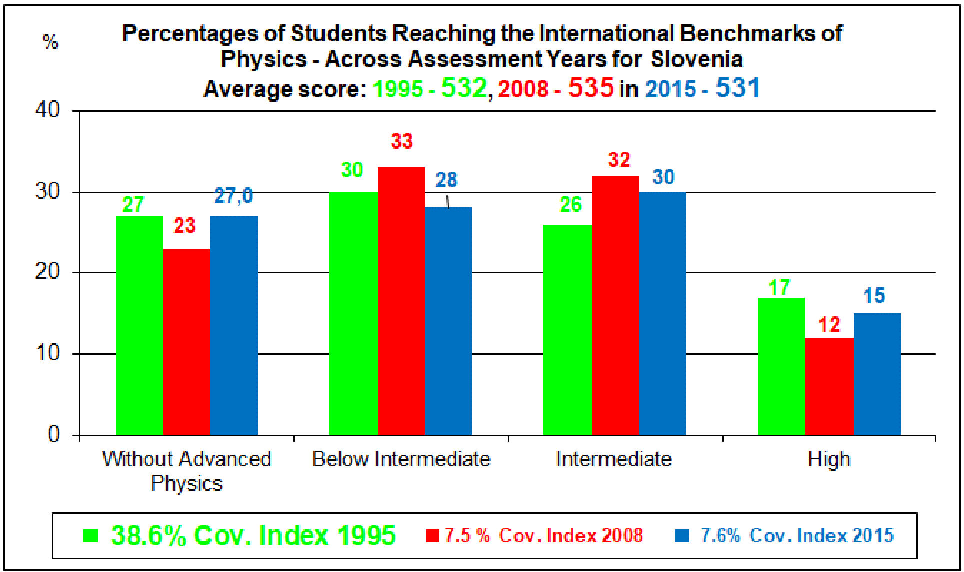
High Benchmark
Students can apply a broad range of mathematical concepts and procedures in algebra, calculus, geometry, and trigonometry to analyse and solve multi-step problems set in routine and non-routine contexts.
Students can analyse and solve algebra problems, including problems set in a practical context. They can solve problems requiring interpretation of information related to functions and graphs of functions. They can determine a sum of an arithmetic sequence and solve quadratic and other inequalities. They can simplify logarithmic expressions and multiply complex numbers.
In calculus, students have a basic understanding of continuity and differentiability. They can analyse equations of functions and graphs of functions. They can relate the graphs of functions to graphs and signs of their first and second derivatives. Students show some conceptual understanding of definite integrals.
Students can use trigonometric properties to solve a variety of problems involving trigonometric functions and geometric figures. They can use the Cartesian plane to solve problems, identify a vector perpendicular to a given vector, and prove that a quadrilateral given in the coordinate system is a parallelogram.

Advanced Benchmark
Students demonstrate thorough understanding of concepts, mastery of procedures, and mathematical reasoning skills. They can solve problems in complex contexts in algebra, calculus, geometry, and trigonometry.
In algebra, students can reason with functions to solve pure mathematical problems. They demonstrate facility with complex numbers and permutations and can find sums of algebraic and infinite geometric series.
In calculus, students demonstrate thorough understanding of continuity and differentiability. They can solve problems about optimization in different contexts and justify their solutions. They can use definite integrals to calculate the area between two curves.
Students use geometric reasoning to solve complex problems. They use properties of vectors to express relationships among vectors. They can use trigonometric properties including the sine and cosine rules to solve non-routine problems about geometric figures.

The graph below shows the difference in TIMSS Advanced Physics achievements across assessment years for Slovenia. The graph also shows percentages of students reaching the international benchmarks of physics achievement across assessment years for Slovenia (IEA, 2016; Exhibit P2.3.)

TIMSS Advanced describes achievements at three International Benchmarks along the scale: **Advanced**, **High** and **Intermediate**. The difference of the coverage index to 100 % present the proportion of those **Without advanced physics**.

The TIMSS Advanced Physics achievement across assessment years for Slovenia are stable.
Almost the same average physics achievement scores (532) have been reached by 7.6 % of the generation in 2015; however, twenty years prior, almost the same average scores (531) have been reached by 38.6% of the generation!



Students Struggle to Reach the TIMSS Advanced International Benchmarks.
TIMSS Advanced describes achievement at three International Benchmarks along the scale: **Advanced**, **High**, and **Intermediate**. There are descriptions of performance at the International benchmarks of physics achievement (IEA, 2016; Exhibit P2.2: for 2015)

Intermediate Benchmark
Students demonstrate some basic knowledge of the physics underlying a range of phenomena. They use their knowledge of forces and motion to solve problems, apply knowledge of heat and temperature to energy transfers, and of conservation laws to everyday and abstract contexts. They show knowledge of electric fields, point charges, and electromagnetic induction. Students apply knowledge of phenomena related to mechanical and electromagnetic waves and knowledge of atomic and nuclear physics to solve problems. Students interpret information in diagrams and graphs to solve problems, calculate a variety of physical quantities in a range of contexts, and evaluate statements to identify explanations for physical phenomena.

High Benchmark
Students apply basic laws of physics in solving problems in a variety of situations. They apply knowledge of forces and motion, communicate understanding of the laws of conservation of energy and momentum, and apply knowledge of heat and temperature to solve problems. Students apply knowledge of Ohm's Law and Joule's Law to electric circuits, solve problems involving charged particles in magnetic fields, and apply knowledge of magnetic fields and electromagnetic induction to solve problems. They show understanding of phenomena related to electromagnetic waves and knowledge of nuclear reactions. Students interpret information in complex diagrams and graphs depicting abstract concepts, derive formulas and provide calculations of a variety of physical quantities in a range of contexts, evaluate explanations for physical phenomena, and provide brief explanations to communicate scientific knowledge.

Advanced Benchmark
Students communicate their understanding of laws of physics to solve problems in practical and abstract contexts.
They apply knowledge of the motion of objects in freefall, of heat and temperature, and of electric circuits and electric fields. Students communicate understanding of magnetic fields and of phenomena related to mechanical and electromagnetic waves, and demonstrate understanding of atomic and nuclear physics.
Students design experimental procedures and interpret results, synthesize information in complex diagrams and graphs depicting abstract physics concepts to solve problems, provide multi-step calculations of a variety of physical quantities in a range of contexts, draw conclusions about physical phenomena, and provide explanations to communicate scientific knowledge.

Good results at the end of pre-university education in Slovenia in Advanced Mathematics and Physics (compared to other 9 countries) can most likely be attributed to positive feedback of the Rules on admission to higher education as well.

However, considering the results of our students 7 or 20 years ago, we can see that students at that time achieved equal (Physics) or even higher results (Mathematics) in a much larger proportion of population! The proportion of population in Mathematics was twice as large (75.4% compared to 34.4%); with Physics the proportion was 5 times larger (38.6% in 1995 compared to 7.6% in 2015).

What has happened since? In Slovenia, education was infused with early choice, as this was seen as friendlier and more modern. In Slovenia, but perhaps in other countries even more so, this resulted in the avoidance of hard work and demanding content.

Today's society glorifies choice, as well as the idea that choice is always in the people's best interest (Renata Salecl, *The Tyranny of Choice*). This is obviously not true. Life choices, including education, are too often described in the same terms as consumer choices: we set out to select the optimal school subject or school program in the same way as we set out to find the right kind of hair conditioner.

By employing 'modern' approaches, we would like to protect students from any effort, stress, discomfort, engagement and duties. We want to give them choice, an individual approach, comfort; we support their ego(-ism). We idolize their opinion (even if irrelevant); anything they say or do is OK, they can always change their minds without consequences, things can be always changed, the present moment is the only important thing, and diligence is not necessary.

Families with low socioeconomic or cultural background, groups on the fringe of society, are especially vulnerable under such unstable, continually changing and unclear circumstances.

We should recognize that by introducing choice in education (too) early, we hinder the long-term interests of individuals and society. We tend to offer choice of: educational tracks, school subjects, levels of difficulty, educational approach, etc. It is irresponsible for policymakers in education and teachers to shift the burden of decisions about students' educational careers (like the level of difficulty or choice of school subjects required for admission to a certain study program, or to successfully work in a certain profession – doctor, teacher, engineer, psychologist) onto underage individuals with little or no knowledge.

20 years of history on TIMSS Advanced results in Slovenia is the argument that we should have (Andreas **Schleicher**, PISA, OECD):

- high universal educational standards for all;
- incentives for students to take tough courses and study hard;
- have clear ambitious goals aligned with high stakes gateways at each stage of their education; and
- use achievement data over time to make decisions.

The case message should be – to paraphrase Obama's speech: "Yes we did. Yes we can - what we have to!"

XIII. Executive summary

Program for the International Assessment of Adult Competencies (PIAAC)

Countries greatly differ in literacy and/or numeracy proficiency among adults (PIAAC).

PIAAC data shows that 25-65 year-old Italian adults with tertiary education achieve similar literacy proficiency (281.3) than Japanese adults without tertiary education (281.1). Japanese adults with upper secondary education are more proficient in literacy (286.7) than Slovenian adults with tertiary education (285.7). Dutch adults with upper secondary education are more proficient in literacy (283.1) than Spanish adults with tertiary education (282.0). Dutch adults with upper secondary education are equally proficient in numeracy (280.7) as Italian adults with tertiary education (280.2). Danish adults without upper secondary education are equally proficient in numeracy (241.0) as USA adults with upper secondary education (241.0). Norwegian adults without upper secondary education are equally proficient in literacy (251.4) as Slovenian adults with upper secondary education (251.6).

There is no easily recognizable pattern of association between proficiency and proportion of students with tertiary education. In other words, we cannot link high proficiency results (quality) with either big or small proportion of population with tertiary education (quantity).

Long-lasting effects

Gustafsson (2015) tested if the effects of education stay with a person through their lifetime into adulthood through analyses of the Swedish PIAAC and PISA data (at the age of 15). Performance differences between age-groups were compared with results obtained for corresponding samples in the five PISA studies and in the IEA studies (TIMSS, PIRLS…) conducted after 1990.

PISA achievement trends at age 15 are reflected in differences in the level of performance between adjacent age-groups in PIAAC. Average PIAAC scores for young adults in Sweden are lower, indicating that they were not able to compensate for deficiencies in knowledge at compulsory school through increased learning (measured in PISA from 2003 to 2012) at a later stage.

The level of achievement at the end of compulsory school thus lasts up to at least age 27, perhaps even life-long (Gustafsson 2016a, 2016b & Rosdahl, 2016). Also, the Sulkunen and Malin (2017) study for Nordic countries shows that the significance of the length and scope of the initial education in developing literacy proficiency overall is difficult to compensate.

The role of basic education cannot be ignored even after completing secondary and tertiary degrees! What is happening within the country in pre-tertiary education is crucial for the quality in higher education as well.

Quantity & Tertiary Education

Around the world, countries have been pushing to expand education – especially at the tertiary level. The policy discussion often appears to assume that the skills of college graduates are exogenous and fixed, implying that the expansion of higher education will lead to proportionate increases in knowledge capital (Hanushek, 2016).

With expanding tertiary schooling, the skills of admitted students are clearly below the average skills before expansion, and expansion of students attending tertiary education is likely to lower the skills of the average graduate. The skills of college graduates (value added) are endogenous and depend directly on skills at entry to college.

The output of higher education – the skills of college graduates – depend directly on admissions standards for colleges and universities. When admission standards are high and demanding, skills of those applying for admission will increase!

Economic growth (GDP) is highly related to the knowledge capital of the country

Economic growth is highly related to the knowledge capital of the country (Hanushek, 2016). The association between years of schooling and growth turns insignificant and its marginal effect is reduced to close to zero once cognitive skills are included in the model. After controlling for the initial level of GDP per capita and for years of schooling, the test score measure features a statistically significant effect on the growth in real GDP per capita in 1960 – 2000 (Hanushek and Wößmann, 2008).

Test scores that are larger by one standard deviation (measured at the student level across all OECD countries in PISA) are associated with an average annual growth rate in GDP per capita that is two percentage points higher over the whole forty-year period (Hanushek and Wößmann, 2008).

School attainment has no independent effect over and above its impact on cognitive skills. There is a big difference between the "quality approach" and the "access approach". School policy can, if effective, raise cognitive skills (Hanushek & Wößmann, 2012). Policy-makers interested in advancing future prosperity should focus on educational outcomes, rather than just inputs or length of study. Similar measures appear promising for the European system of higher education (Wößmann, 2014).

Achieved cognitive skills and proficiency in literacy skills impact non-educational and non-economic features as well

Knowledge and skills of the population are a leading determinant of economic growth, employment, and earnings in modern knowledge-based economies such as the EU (Wößmann, 2014). Education reduces risk of unemployment.

Workers with higher PIAAC proficiency in information processing skills are more likely to be employed (in some countries), earn higher wages (in most countries) and have better social outcomes (in all countries).

Across countries and economies, there is a positive correlation between skills proficiency in literacy and trust, volunteering and political efficacy (with correlation coefficients in the order of 0.40).

Research findings and some lessons from best performers

Selected lessons from best performers (Andreas Schleicher, PISA, OECD)

- Commitment to universal achievement; **universal educational standards** and personalization as the approach to heterogeneity in the student body (as opposed to a belief that students have different destinations to be met with different expectations and selection/stratification as the approach to heterogeneity).
- A **well-established delivery chain** through which curricular goals translate into instructional systems, instructional practices and student learning (**intended, implemented and achieved curriculum**).
- Clear articulation of **who is responsible for ensuring student success and to whom**.
- Aligned system incentive for students: **students have incentives to take tough courses and study hard**.
- Have clear and ambitious **goals** that are shared across the system and **aligned with high-stakes gateways and instructional systems**.
- **Gateways affect** the strength, direction, clarity and nature of the incentives **influencing students at each stage of their education.**
- Provide **a data-rich school environment** to combat inequities. "*Without data, you are just another person with an opinion.*" (William Edwards Deming)
- The results from PISA show higher levels of socio-economic equity in school systems that **use achievement data to make decisions about the curriculum and track achievement data over time** (OECD 2010, 47).

Many traditional policies of simply providing more funds for schools or adding specific resources such as smaller classes do not provide much hope for significant improvements in student achievement. Policy makers should focus on the outcomes for individuals in terms of skills not just the inputs as the level of education in terms of years of schooling (Hanushek, 2016).

A substantial part of the cross-country variation in student achievement corresponds with differences in **external exams**, **school autonomy**, private competition and **early tracking** (Wößmann, 2016).

External exams

Research results (Jürges, Shneider, Büchel 2003) suggest that centralized examinations increase student achievement by about one third school year equivalent. Positive effects of external exit examinations on test-score outcomes have been shown for Canadian provinces, for US states and for German states. International cross-country evidence suggests that the effect of external exit examinations on student achievement may well exceed a whole grade-level equivalent (Piopiunik, Schwerdt, Wößmann 2012).

Internal grades in Slovenia are inflated (Zupanc, Bren, 2010), meaning that higher grades do not derive from a corresponding increase in student achievement. Sweden is also the case; it has a big problem with inflation of internal grades. The average school-leaving grade rises almost continuously year per year, while the opposite applies to the PISA scores (Henrekson, & Jävervall, 2016).

Teachers in Slovenia give the spectrum of lower and higher grades regardless of measured achievements relative to absolute (criterion) reference. As a rule, high achievements for the external parts is followed by low achievements for the internal parts and vice-versa; low achievements for the external parts is increased by high achievements for the internal parts. The criterion is higher for „better" schools and lower for schools with „lower" achievement students.

According to literature, internal grading is biased also with regard to Socio-Economic Status (SES), gender, etc. Teachers with internal grading tend to give advantaged SES students and girls better school grades, even if they do not have better performance and attitudes than disadvantaged SES students and boys (OECD, 2012).

School autonomy and accountability

The relationship between autonomy, accountability and performance is complex. It is a combination of several autonomy and accountability policies, not just a single, isolated policy relating to better student outcomes (OECD, 2011). PISA results suggest that, when autonomy and accountability are combined in an intelligent manner, they tend to be associated with better student performance (OECD, 2011).

Across OECD countries, those that use Curriculum-Based External Examination Systems (CBEES) tend to perform higher, even when accounting for national income: students in school systems that use CBEES perform, on average across OECD countries, 16 points higher than students in school systems that do not use these examinations (OECD 2010). Greater autonomy in decisions relating to curricula, assessments and resource allocation tends to be associated with better student performance, particularly when schools operate within a culture of accountability. There is an indication that local decision-making (autonomy) works better when there is also external accountability that limits any opportunistic behaviour of schools.

Equity in education

Equitable systems ensure that also the outcomes of education and training are independent of socio-economic background (SES) and other factors that lead to educational disadvantage, and that treatment reflects individuals' specific learning needs" (EU, 2006). All countries should try to meet two goals to ensure the best outcome for their schools: getting high levels of student achievement (quality) while minimizing systematic gaps in performance - equity (Wößmann, 2016). Poor children have a much higher risk of low school achievement (OECD, 2012).

In countries with earlier tracking, the achievement difference between children with different socio-economic backgrounds is considerably larger (Hanushek, Wößmann, 2011). Early tracking has especially negative effects on the achievement levels of disadvantaged children (COM; 2016).

Disadvantaged students hold lower expectations of further education, even when they perform as well as students from advantaged backgrounds. Inequalities in expectations are present in practically all countries and economies. Differences in expectations between students who have similar performance levels but who come from either advantaged or disadvantaged backgrounds is highest in Korea, Latvia, Poland, Serbia and Slovenia. In all these countries, socio-economically advantaged students are four times less likely to expect to end their education after completing upper secondary school than disadvantaged students who perform equally well (OECD, 2012, 35-36)

Entering different tracks of upper secondary education in Slovenia is highly associated with students and their families' socio-economic status (SES).

A common assumption has been that a "free" system of higher education (one funded entirely by the state) is equitable by itself. In reality, this assumption did not materialize, since the main determining factor in participation is socioeconomic background. Free access to higher education does not necessarily guarantee equity (COM; 2006). The growth in the number of students did not increase equity, as it mostly favoured individuals from higher socio-economic groups or those whose parents with higher education.

Data-rich school environment

Within education systems, a culture of evaluation is needed. Effective long-term policies must be based on solid evidence. As data based decision-making is receiving increased attention in education, more and more school performance feedback systems (SPFS) are being developed and used worldwide. These systems provide schools with data on their functioning (Verhaeghe, Schildkamp, Luyten, Valcke, 2015).

The Slovenian National Examinations Centre (NEC) also developed a version of SPFS for purposes of ensuring effectiveness and attaining improvement in classrooms and schools in upper secondary education in Slovenia: Assessment of/for Learning Analytic Tool (ALAT) (Zupanc, Urank, Bren, 2009). Software enables interactive analysis of data for each school with national benchmarks, calculated using the same selection criteria (Brejc, Sardoč & Zupanc, 2011).

The NEC has also results of the National Assessment (NA) at the end of compulsory education – in the last year (9th grade) of basic school. There are data for achievement of Mother Tongue, Math and 1 optional subject for each student in the generation in Slovenia. Value-added approach enables schools to analyse the progress of their students from the end of basic school to the end of upper secondary school.In 2014, the software was upgraded with new features: per item analysis for criterion-referenced interpretation and value added analysis.

Teachers, head teachers and other educators in Slovenia are able to access the examination database, analyse assessment results and teachers' grades, interpret the achievement of their students, and analyse the efficiency of teaching and learning in classrooms and schools in pre-university upper secondary schools with the General and Vocational Matura as the exit examinations.

Caution due to misuse of data

While sound governance calls for use of data when making decisions we must ensure high standards for data quality and interpretation of it. Education is an important subsystem in every country and decisions should be informed, evaluated and (later) reevaluated. Data is in this context very important as it provides arguments for or against certain decisions and can help us – but only if interpreted correctly! What follows are three real case examples that show how we can reach wrong conclusions and only after careful examination and interpretation of data can we make better judgement.

The Programme for International Student Assessment (PISA) 2006 raised many questions when Slovenia got very low results on the equity indicator in PISA 2006. Slovenia ranked third highest on the between-school variance indicator – an indicator of great differences between schools. It seemed in Slovenia students were attending schools that (over time) created remarkable differences between students, which leads to the conclusion that some students were in 'good' schools and some students in 'bad' ones and the educational system was promoting this. Only careful reexamination of data demonstrated the error in interpretation.

In Slovenia, upper secondary schools mainly fall into two major tracks – schools that end in General Matura and lead to universities, and schools that end in Vocational Matura and provide vocational qualifications. However, students in programmes ending in Vocational Matura can apply for tertiary education as well, and since there's no established measure how different examinations in General and Vocational Matura actually are in terms of difficulty they are considered equal in admission process! Examination of circumstantial data demonstrated this most probably isn't correct.

Third example demonstrated a situation in statistics, referred to as Simpson's paradox. When we look at science achievement at PISA 2006 in Slovenia in each educational track individually, boys outperform girls in their science achievements by 10, 20 or even more than 30 points. These differences are statistically and practically significant. Naturally, you would expect that boys on the whole outperform girls in science achievements. However, the results negate this conclusion – girls outperform boys by 7 points! There is nothing wrong with the student sample, calculations made or statistics used. The results are the consequence of differences in proportions of boys and girls in those educational tracks. Again it is a lesson that data should be interpreted with care.

Pre-university TIMSS Advanced study

Passing the Matura exam with compulsory math proves general ability for any academic course. The math score is even more important for students applying to the most prestigious faculties that have enrolment thresholds. According to the Rules on admission to higher education, these students should take math at the most advanced level, and many of them should prove excellence by taking physics or chemistry or biology or two of them4 as additional Matura subjects. Good results at the end of pre-university education in Slovenia in TIMSS Advanced Mathematics and Physics (compared to other 9 countries) can most likely be attributed to positive feedback of the Rules on admission to HE as well.

However, considering the results of our students 7 or 20 years ago, we can see that students at that time achieved equal (Physics) or even higher results (Mathematics) in a much larger proportion of population! The proportion of population in Mathematics was twice as large (75.4% compared to 34.4%); with Physics the proportion was 5 times larger (38.6% in 1995 compared to 7.6% in 2015).

What has happened since? In Slovenia, education was infused with early choice, as this was seen as friendlier and more modern. In Slovenia, but perhaps in other countries even more so, this resulted in the avoidance of hard work and demanding content.

20 years of history on TIMSS Advanced results in Slovenia is the argument that we should have (Andreas Schleicher, PISA, OECD):

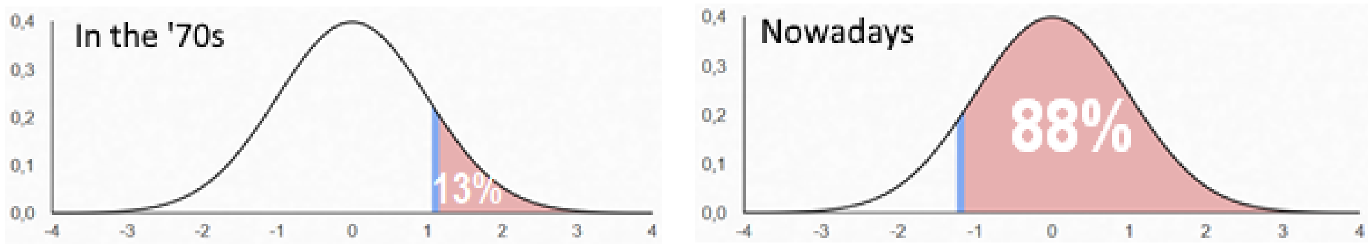
- high universal educational standards for all;
- incentives for students to take tough courses and study hard;
- have clear ambitious goals aligned with high stakes gateways at each stage of their education; and
- use achievement data over time to make decisions.

The case message for Slovenia should be, to paraphrase Obama's speech: "*Yes we did. Yes we can - what we have to!*"

XIV. Conclusions

The main issue with admission to tertiary education in Slovenia is not the use of the Matura (external examinations) vs. admission exams, but the quantity (proportion) of admitted students and the quality of their knowledge.

<u>Neither the Matura examination nor motivational letters from candidates, interviews or admission exams can provide all universities and vocational schools with excellent freshmen when such large proportions of population are admitted.</u>



Graph X: Admitted in tertiary education in Slovenia (proportion of those with upper secondary attainment in the yearly cohort).

The issue of quality and equity of education is associated with all levels of education and works in both directions at every transition point to a higher level. **The effect is mutual.**

If the admission process sets high expectations and is very demanding in terms of knowledge, students will demonstrate higher knowledge at the end of high school and at the Matura examinations.

Due to traditions in national curricula and Slovenia's small population, we have excellent and centrally available data (even on NEC) about students' achievements before they enter universities – we just have to use it properly.

The impression that lowering expectations helps underprivileged parts of the population and therefore increases equity **is wrong**! Lowering of standards, inflation of school grades, certificates and diplomas (all of which hinder the quality of education) do not increase equity.

A less demanding school is just another way of saying that children from the lower middle class are placed in a circle of people without access to prestigious positions in society (Gaber & Tašner, 2009).

Quality **or** equity is therefore not the right question; our goal should be quality **AND** equity!

When the **educational system** promotes and rewards students by lowering standards on all levels, it promotes **lower knowledge**. This turns into a systemic issue, as the entire chain (schools) is only as strong (or weak) as its weakest link (graduates).

There are no straightforward solutions. For Slovenia, with a very large proportion of annual population admitted to tertiary education, one solution seems to head in the direction of smaller proportions and higher demands, even if this leads to fewer students enrolled, fewer certificates, diplomas, PhDs awarded. We should enable universities to keep high standards and stop the vicious spiral of lowering the demands for students. This should result in higher quality.

A culture of evaluation is needed within education systems. Effective long-term policies must be based on solid evidence. For Member States to fully understand and monitor what is happening in their systems, they need channels for producing and accessing relevant research, a statistical infrastructure capable of collecting the necessary data, and mechanisms to assess progress as policies are implemented.

Member States should develop a culture of evaluation. They should develop policies for the entire lifelong learning continuum which take full account of effectiveness and equity in combination and in the long term, and which complement policies in related fields (COM, 2006).