CHAPTER 7

The Role of International Benchmarking in the Convergence/Divergence of European Education

Maja Mihaljević Kosor, Jurica Pavičić, and Nikša Alfirević

Abstract In this chapter, the authors analyse and discuss conceptual and methodological issues related to international benchmarking and its effect on educational policy and practice in selected European countries. Benchmarking is viewed as a process of comparing educational outcomes in schooling systems, identifying strengths and weaknesses and serving as a platform to identify best-practice examples, thus providing a foundation

M. Mihaljević Kosor (✉)
Faculty of Economics, Business and Tourism, University of Split, Split, Croatia
e-mail: majam@efst.hr

J. Pavičić
Department of Marketing, Faculty of Economics and Business, University of Zagreb, Zagreb, Croatia

N. Alfirević
Department of Management, Faculty of Economics, Business and Tourism, University of Split, Split, Croatia

L. Moos et al. (eds.), *Educational Leadership, Improvement and Change*, Palgrave Studies on Leadership and Learning in Teacher Education, https://doi.org/10.1007/978-3-030-47020-3_7

for school development and other improvements in the education system. Using data from international surveys, the authors compare educational outcomes over time for a sample of European countries, discuss recent trends and identify the main similarities/differences. The data on educational outcomes come primarily from the wide-ranging OECD Programme for International Student Assessment (PISA), as well as from other relevant sources of data on European education. The chapter is concluded with the lessons learnt from international benchmarking, in the context of challenges in the social and education environment of the selected European countries.

Keywords International benchmarking • PISA (Programme for International Student Assessment) • OECD (Organisation for Economic Co-operation and Development) • Best practices

Introduction

There is widespread recognition of the importance of education and skills in the formation of human capital and increasing pressure to demonstrate that the sizeable public and private funding of education pays back economically, socially and culturally (OECD, 2017a). A substantial amount of research has been carried out focusing on the inputs and outputs of the education system, in most cases in order to obtain some ranking of educational institutions in terms of efficiency scores. In this regard, benchmarking is viewed as an extremely important tool in improving the efficiency of the education sector.

Key Concepts and Challenges for Benchmarking in Education

In the education sector, *economic efficiency* can be viewed in terms of alleviating the costs of the public sector, given that most of the primary and secondary schools are funded from local/regional or national budgets. Efficiency refers to the degree to which the provision of goods/services is maximised with the available resources. The commonly examined types of efficiency in education are technical and allocative efficiency. Blank (2000)

defines technical efficiency as the extent to which services can be expanded without increasing resources, or, alternatively, the amount by which resources can be contracted without reducing the services. Hoxby (1995) argues that allocative efficiency is 'getting the amount of education right' (p. 54), while productive efficiency concerns minimising the cost of the schooling provision. *Effectiveness*, on the other hand, ascertains whether a specific set of educational resources contributes to student performance and to what extent. As Drucker (1967) states, efficiency is about 'doing things right', while effectiveness is about 'doing the right things'. Nevertheless, education performance depends on the interpretation of multiple stakeholders (students, parents, teachers, school principals, the local community and government), functioning in a network of principal-agent relationships and having different, sometimes even conflicting, objectives. In addition, educational institutions use multiple inputs to produce a range of outputs, making it difficult to assess what contribution to output is made by the education processes and/or by the enrolment of the 'right' students.

Education policy needs to improve the outcomes of education, while balancing the efficiency and equity of the system. In debates over the equity-efficiency trade-off, it is often pointed out that the pursuit of greater equity usually leads to a decrease in efficiency, and vice-versa. For example, in UK secondary schools, the goal of increasing efficiency may result in some form of stratification of poorer students in schools with the weakest performance, and consequentially the resources get allocated to better-performing schools (in Bradley, Johnes, & Millington, 2001). There are also some dysfunctional effects in the education sector where schools may be trying to improve their position in the league tables through 'cream-skimming' or grade inflation (Adnett, Bougheas, & Davies, 2002). As Hanushek (1997) notes, in the economic analysis of education, the focus has shifted from the quantitative analysis of inputs in education to the measurement and analysis of student performance and outcomes.

International benchmarking of education outcomes, especially workplace-related skills (Hopkins, Pennock, Ritzen, Ahtaridou, & Zimmer, 2008), creates a new equalising force in the policy landscape. This is especially significant for established international benchmarking studies, such as the Programme for International Student Assessment (PISA), which is developed and implemented by the Organisation for Economic Co-operation and Development (OECD). It is a global

triennial survey of student performance, testing the problem-solving skills and cognition of 15-year-olds.

There is plenty of evidence of the impact of PISA on the convergence of national education policies. National responses to the release of PISA results have been diverse (Martens, Nagel, Windzio, & Weymann, 2010). Out of countries included in this study (Switzerland, Germany, New Zealand and the USA), both Switzerland and Germany prompted significant education reforms in light of perceived poor performance. This is in contrast to the United Kingdom's lack of policy response to moderate performance, which Knodel and Walkenhorst (2010) attribute to education reforms already underway in England in the years prior to the release of the PISA results. The authors of another study asked members of the PISA governing boards to comment on the effect of PISA scores on national education policies (Breakspear, 2012). Representatives from 17 countries (out of 37 that participated in the study) found PISA to be 'very influential' in informing their policy-makers, and a further three countries found PISA to be 'extremely influential' (Denmark, England and Japan). The respondents also indicated that PISA scores prompted changes in 19 countries and led to partial changes in a further 11 countries. The use of PISA data has also been criticised, even within Finland, as a country scoring highly on the PISA list(s). Rautalin and Alasuutari (2009) argue that the Finnish central government used the PISA data and reports without much critical scrutiny, attributing success in the Finnish students' rankings to their own actions, as well as blaming other actors for the shortcomings of the education system.

Other critical voices argue that the PISA methodology is not robust (Kreiner & Christensen, 2014), although there is some empirical evidence speaking in favour of the existing methodology (Jerrim et al., 2018). Some studies argue that test items might be biased, either due to translations and/or cultural sensitivities, leading to the variability of test items in different countries (Goldstein, 2017). Even if those items were removed, or reduced by improved methodology, a critical argument views that the PISA rankings provide, essentially, an extremely simplified 'black box'. The policy community should trust that these scores actually capture the complexity of education systems and synthesise the quality of a country's schooling, without complete transparency over the process of their construction and implementation. It is also noteworthy that additional country- and school-level data (coming from, e.g. PISA questionnaires for principals and students) seem to be largely ignored by the policy actors,

while some significant policy decisions, such as school funding and education staff salary increases, could be based on superficial readings of the PISA national table rankings (Sjöberg, 2016).

In addition, *international benchmarking purposefully puts the emphasis on the reading, science, mathematics and, recently introduced, financial literacy related skills, instead of traditional school curricula, that is, knowledge domains.* It could be argued that such a trans-national approach has distinct advantages in terms of transcending the traditional limitations in measuring the critical thinking and creativity skills relevant for students' future success in the contemporary world (Schleicher, 2018). On the other hand, such a methodology sends not-so-subtle signals to national educational policies regarding the importance of individual knowledge domains (Sjöberg, 2019). PISA results also reflect the institutional and political power of governments involved in the functioning of the OECD (ibid.), as opposed to alternative benchmarking studies, such as *TIMMS (Trends in International Mathematics and Science Study)* or *PIRLS (Progress in International Reading Literacy Study)*, which are designed and implemented by the less 'prestigious' *IEA (International Association for the Evaluation of Educational Achievement)*.

The outcome of such a global trend in education policy convergence is most certainly compatible with the developments of directing education away from the notion of *Bildung* and ideals of democracy and solidarity towards standardised education relevant for the development of human capital (as prescribed by the labour market), but also susceptible to standardised measurement and management practices (Meyer & Benavot, 2013).

Comparing the Educational Performance of Selected European Countries

The goal of this section is to provide a comparative analysis of the educational performance of selected European countries. The data used in this analysis come from the three recent PISA rounds (conducted in 2012, 2015 and 2018). PISA data are combined with country specific data to gain more insight into the main similarities/differences in educational performance between countries over time. The strongest examples are then identified and a closer analysis of their progress in education is discussed.

In general, the PISA data are used to inform policy-makers and other education stakeholders of the efficiency, equity and the quality of education systems. Given that it is one of the most comprehensive studies of educational outcomes in the world, over the years performance in PISA has been used in different types of research, ranging from analysis of its effects on adult life outcomes, to equity in education, economic performance of countries and so on. For example, in terms of its effect on adult life outcomes and equity, a clear link was established: students who did well in PISA tests did better in the labour market or in further education (OECD, 2018). The study also found that students in Denmark who expected to work in a high-skilled job were about 40 percentage points more likely to be doing so as young adults than students who did not have such expectations (and about 10 percentage points in Switzerland). Data for this analysis are only available for four countries (Australia, Canada, Denmark and Switzerland) for students who took the test in 2000 and 2003 and were followed ten years later in their transition to early adulthood. In this case, PISA results were good predictors of success in higher education and in the job market later on. An important study linking PISA scores and other economic and education indicators found that an increase of 25 points in average PISA scores in the next 20 years would result in a combined gain for OECD countries to the amount of USD 115 trillion for the generation born in 2010 (OECD, 2010).

In the following sections, we examine the trends in the educational performance of European countries in order to establish common features and identify countries which have made rapid improvements in educational performance over recent years. PISA data are complemented with data on government expenditures on education, the number of students and teachers in secondary education, and teacher salaries. The data are for the EU-28 countries, with special focus on countries appearing in country reports and specific analyses in the earlier chapters, that is, Croatia, Slovenia, Denmark, Lithuania and Norway.

The PISA data for the EU-28 in 2012, 2015 and 2018 point to Estonia, Finland, Ireland, Poland, Sweden, the United Kingdom and Denmark as the best-performing countries, with a reading score above 500 points (the international average in reading for OECD countries is 487 points). Bulgaria, Cyprus, Romania and Malta are the worst-performing of the EU-28 countries, with a score under 450 points (in Fig. 7.1). The situation is similar in other tested areas (maths and science), where Bulgaria, Romania and Greece are the worst-performing countries. This is not

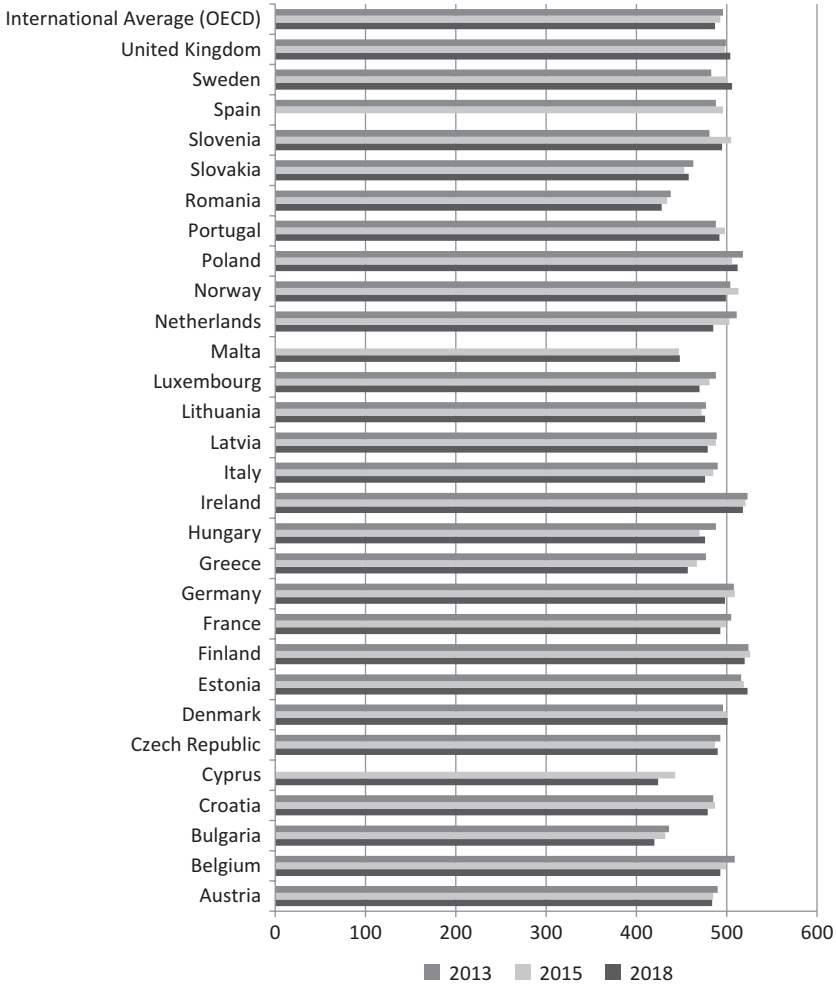

Fig. 7.1 Average PISA scores in Reading in 2012, 2015 and 2018 for the EU-28 and Norway. (Source: Authors, based on OECD data (https://www.oecd.org/pisa/data/). Note: The countries are ordered from high to low scores in 2018. There are no 2018 data for Spain or 2012 data for Cyprus and Malta)

surprising, given that countries with better PISA scores are those with higher GDP per capita. The emphasis here was on performance in reading, given the focus of the last PISA round.

Nevertheless, the OECD's analysis of Skills for Jobs found that there are critical shortages in maths, science and problem-solving skills for its member countries and for the occupations related to them, such as science, engineering and the Information and Communications Technology (ICT) professions (OECD, 2017b). Considering the average score in all three tested areas in 2015, most of European 15-year-olds underperform when compared to their peers in Asian countries (European Commission, 2018). Furthermore, nearly 20% of students in OECD countries do not attain the baseline level of reading proficiency.

The situation presented in Fig. 7.1 raises an alarming issue, which points to a decrease in average performance across the majority of countries. Both OECD and EU-28 average scores decreased in 2018 in comparison to previous years. In comparison to 2012, the EU-28 average in reading in 2018 dropped by ten points and by nine points in OECD countries. The decrease in average maths scores for the EU-28 was from 491 to 489 points and from 494 to 489 for OECD countries. In science, the average score fell from 499 to 484 for the EU-28 and from 501 to 489 for OECD countries. It may be concluded that there was no real improvement in learning outcomes of students in the EU-28 and OECD countries, although expenditure on schooling rose by 15% over the past decade (OECD, 2019a).

We next turn to countries presented in more detail in this volume. We look at their mean performance in reading over time and investigate the direction and trajectory in their mean performance (in Fig. 7.2).

Croatia, Slovenia and Lithuania started participating in PISA in 2006; hence there are no previous data for these countries. Only Denmark and Lithuania had an increase in their average score in 2018 when compared to the 2015 assessment, while other countries had a noticeable drop in their average reading score. In terms of the overall direction of the trend (the sign and significance of the average three-year trends), all countries had no significant average trend. Only Slovenia had a U-shaped trend, which was more positive over more recent years.

OECD (2019a) data provide brief snapshots of trends for selected countries. In Croatia, the mean performance in reading and mathematics remained stable, around a flat trend line. The mean performance in science declined, demonstrating a widening gap in learning outcomes

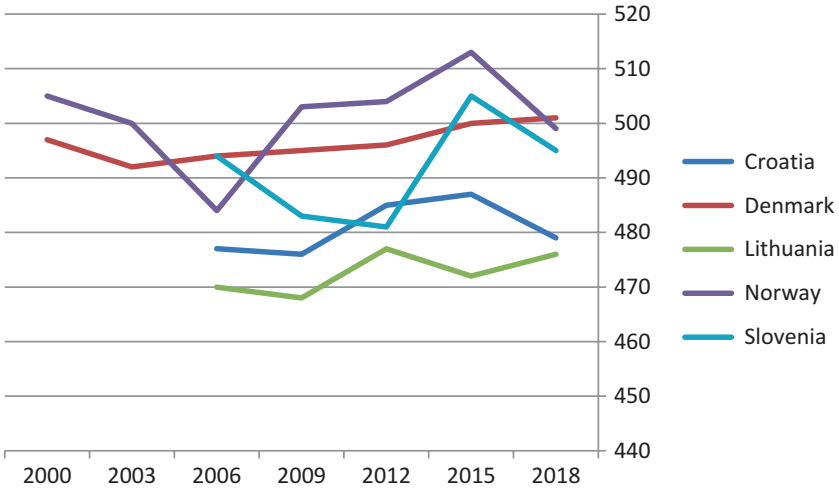

Fig. 7.2 Average PISA score for selected countries across PISA assessments. (Source: Authors, based on OECD data (https://www.oecd.org/pisa/data/))

between the highest- and lowest-achieving students. In Denmark, there was no overall trend in mean performance in reading, maths and science. However, the proportion of top-performing students in reading increased by 3.7 percentage points from 2009 to 2018. Furthermore, there was a narrowing gap in learning outcomes in mathematics between the highest- and lowest-achieving students in Denmark.

In Lithuania, no clear direction or change can be determined for average performance in reading and mathematics. The mean performance in science was found to fluctuate more and suggests a declining trend. In reading, the proportion of top-performing students increased by 2.1 percentage points from 2009 to 2018. Norway's performance in PISA 2018 is below its average performance in reading and science from the 2015 PISA round. When a longer time period is considered, there was no discernible direction or change. The fall in performance for Norway is to a certain extent influenced by the increase of the population of immigrant students from 2009 to 2018. These students tend to score below native-born peers.

In Slovenia, the overall performance trend in reading and mathematics is U-shaped and has been more positive over recent years, while the trend

in science is steadily negative. The performance in science was lower, on average, than in 2006 and 2015. There is an improvement in reading performance concentrated among the highest-achieving students, and the proportion of top-performing students in reading increased by 3.2 percentage points from 2009 to 2018. In contrast, the proportion of top-performing students in science decreased by 5.6 percentage points in the same time period.

Across the OECD countries, only 77% of students attained Level 2 proficiency in reading (i.e. students have acquired the technical skill to read and can use reading for learning). Only Lithuania is below this average, while Croatia, Norway, Slovenia and Denmark (sorted in ascending order) are above. Girls outperformed boys in reading in the PISA 2018 assessment, while the proportion of top performers between boys and girls increased in Slovenia and Denmark. Slovenia and Denmark also exhibited a narrowing of the gender gap in mathematics performance.

The impact of socio-economic background on average PISA scores greatly varies across countries. For example, Denmark and Norway had students with higher reading performance than the OECD average. At the same time, the relationship between socio-economic status and average reading performance in these two countries was weaker than the OECD average. Some of these differences in performance are also related to the socio-economic contexts of schools, that is, the level of social and skills stratification between schools, where disadvantaged students are more likely to attend schools with higher proportions of low-achieving students. A higher segregation of disadvantaged students is found in Lithuania, while Norway, Croatia, Denmark and Slovenia are below the OECD average. Countries also differ related to performance variation among schools. For instance, Finland has the lowest between-school variation in reading performance (of about 7%), suggesting that the closest school is always the 'best' one, while the Netherlands, Lebanon and Israel had the highest between-school variation (of about 60% and higher). The OECD average for between-school variation is 29%, and Lithuania and Croatia are just above this average (about 32%). Norway and Denmark are well below (around 10%), while Slovenia is above the average at around 40%.

In recent years, the issue of immigration in the EU has received increasing attention. The difference in average reading score between immigrant and non-immigrant students is 41 points. In most countries, immigrant students are likely to be from a disadvantaged socio-economic position, with the largest proportions in Denmark, Norway and Slovenia (along

with Austria, Finland, France, Germany and Greece). In these countries, more than 45% of immigrant students are disadvantaged. However, the performance of immigrant students varies considerably across countries. In Denmark and Slovenia, the share of immigrant students who were in the top quarter of educational performance was as large as the share of disadvantaged students who achieved that level. In some countries, like Australia, Jordan, Singapore and Saudi Arabia, immigrant students scored higher than non-immigrants. Large variations in the average performance of immigrant students remain, even after accounting for their socio-economic status and country of origin. This suggests that there may be a role for educational policy to minimise these differences.

Interestingly, countries with higher schooling expenditures do not necessarily achieve higher average performance. A positive relationship between investment in education and average performance is found, but only up to a threshold of USD 50,000 (of cumulative expenditure for students from the age of 6 to 15) (OECD, 2019b). However, some countries have exhibited significant improvement, even with limited resources. Portugal advanced to the OECD average level, regardless of the financial crisis, while Estonia gained its top position with expenditure per student that is 30% lower than the OECD average (OECD, 2019b).

The data from PISA 2015 indicate that in 34 school systems, particularly in Austria, Belgium, Croatia, France, Germany and the Slovak Republic, students who do not attend regular science lessons are more likely to be in socio-economically disadvantaged schools (OECD, 2019c). A further examination of the role of government funding and performance in PISA (in the 2012 and 2015 rounds) was analysed for a sample of European countries and reported by Mihaljević Kosor, Malešević Perović, and Golem (2019). The authors found a negative relationship between the pupil-teacher ratio and PISA scores. The direction of the influence of government expenditures on education and PISA scores was less clear, given the substantial variations between countries. Furthermore, the authors found that a smaller number of teachers is related to lower PISA scores, and higher teacher salary is related to higher PISA scores.

This analysis can be linked with the Teaching and Learning International Survey (TALIS) which provides more information on teachers, school leaders and the learning environment in schools. In the 2018 survey, data were gathered for 48 countries. In terms of the efficient use of teachers' time, the data from TALIS find only 78% of a typical lesson is dedicated to teaching, while the remainder is the time spent keeping order or dealing

with administration. This represents a decrease in time spent on teaching and learning over the last five to ten years (OECD, 2019c), which could be linked to the education outcomes, as well.

Conclusion

Benchmarking provides incentives for educational systems and stakeholders to increase efficiency, in the absence of competitive pressures. PISA rankings stand out as significant instigators of change in education, although their role as an international benchmarking standard is criticised. Firstly, this relates to the accuracy of the results and rankings. There are wide disparities in PISA results across countries and these differences have not yet been conclusively attributed. Some countries have been found to select only top-performing students from educationally advantageous areas to participate in PISA, thus skewing the results (as reported in Forbes, 2017). Furthermore, a range of policies and activities that have been undertaken in light of PISA often display a low level of policy coherence (noted in Hopkins et al., 2008). Thus, international benchmarking remains an interesting and challenging topic for further analysis and educational policy-making.

References

Adnett, N., Bougheas, S., & Davies, P. (2002). Market-Based Reforms of Public Schooling: Some Unpleasant Dynamics. *Economics of Education Review, 21*(4), 323–330.

Blank, J. L. T. (2000). Public Provision and Performance: Contributions from Efficiency and Productivity Measurement. Amsterdam: Elsevier Science.

Bradley, S., Johnes, G., & Millington, J. (2001). The Effect of Competition on the Efficiency of Secondary Schools in England. *European Journal of Operational Research, 135*, 545–568.

Breakspear, S. (2012). *The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance*. OECD Education Working Papers, No. 71. Paris: OECD Publishing. https://doi.org/10.1787/5k9fdfqffr28-en

Drucker, P. (1967). *The Effective Executive*. London: Heinemann.

European Commission. (2018). Equal Opportunities: Skills, Education and Overcoming Disadvantages. In *Employment and Social Developments in Europe Annual Review 2018*. Luxembourg: Publications Office of the European Union.

Forbes. (2017). Are the PISA Education Results Rigged? Retrieved November 10, 2019, from www.forbes.com/sites/realspin/2017/01/04/are-the-pisa-education-results-rigged/#3b0ff7f81561

Goldstein, H. (2017). Measurement and Evaluation Issues with PISA. In L. Volante (Ed.), *The PISA Effect on Global Educational Governance* (pp. 49–58). Abingdon: Routledge.

Hanushek, E. (1997). Outcomes, Incentives, and Beliefs: Reflections on Analysis of the Economics of Schools. *Educational Evaluation and Policy Analysis, 19*(4), 301–308.

Hopkins, D., Pennock, D., Ritzen, J., Ahtaridou, E., & Zimmer, K. (2008). *External Evaluation of the Policy Impact of PISA*. EDU/PISA/GB(2008)35/REV1. Paris: OECD.

Hoxby, C. (1995). *Is There an Equity-Efficiency Trade-Off in School Finance? Tiebout and a Theory of the Local Public Goods Producer*. NBER Working Paper No. w5265. Cambridge, MA: National Bureau of Economic Research.

Jerrim, J., Parker, P., Choi, A., Chmielewski, A. K., Sälzer, C., & Shure, N. (2018). How Robust Are Cross-Country Comparisons of PISA Scores to the Scaling Model Used? *Educational Measurement: Issues and Practice, 37*(4), 28–39.

Knodel, P., & Walkenhorst, H. (2010). What's England Got to Do with It? British Underestimation of International Initiatives in Education Policy. In K. Martens, N. Alexander Kenneth, M. Windzio, & A. Weymann (Eds.), *Transformation of Education Policy*. Basingstoke: Palgrave.

Kreiner, S., & Christensen, K. B. (2014). Analyses of Model Fit and Robustness. A New Look at the PISA Scaling Model Underlying Ranking of Countries According to Reading Literacy. *Psychometrika, 79*(2), 210–231.

Martens, D. K., Nagel, A.-K., Windzio, M., & Weymann, A. (2010). *Transformation of Education Policy*. Basingstoke: Palgrave.

Meyer, H. D., & Benavot, A. (Eds.). (2013). *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Oxford: Symposium Books.

Mihaljević Kosor, M., Malešević Perović, L., & Golem, S. (2019). The Role of International Benchmarking in Shaping Educational Policy in Small European Countries: Challenges and Implementation Within Europe. In A. H. Ingþórsson, N. Alfirević, J. Pavičić, & D. Vican (Eds.), *Educational Leadership in Policy: Challenges and Implementation Within Europe* (pp. 27–40). Cham: Palgrave Macmillan.

OECD. (2010). *The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes*. Paris: OECD Publishing. https://doi.org/10.1787/9789264077485-en

OECD. (2017a). *Benchmarking Higher Education System Performance: Conceptual Framework and Data, Enhancing Higher Education System Performance*. Paris: OECD Publishing. https://www.oecd.org/education/skills-beyond-school/benchmarking-higher-education-systems-performance.htm

OECD. (2017b). *Getting Skills Right: Skills for Jobs Indicators*. Getting Skills Right. Paris: OECD Publishing. https://doi.org/10.1787/9789264277878-en

OECD. (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/9789264073234-en

OECD. (2019a). *PISA 2018 Results (Volume I): What Students Know and Can Do*. Paris: OECD Publishing. Retrieved December 12, 2019, from https://doi.org/10.1787/5f07c754-en

OECD. (2019b). *PISA 2018: Insights and Interpretations*. Paris: OECD Publishing. Retrieved December 12, 2019, from https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf

OECD. (2019c). *Education GPS*. OECD. Retrieved November 20, 2019, from http://gpseducation.oecd.org

Rautalin, M., & Alasuutari, P. (2009). The Uses of the National PISA Results by Finnish Officials in Central Government. *Journal of Education Policy, 24*(5), 539–556.

Schleicher, A. (2018). *World Class: How to Build a 21st-Century School System, Strong Performers and Successful Reformers in Education*. Paris: OECD Publishing. https://doi.org/10.1787/4789264300002-en

Sjöberg, S. (2016). OECD, PISA, and Globalization: The Influence of the International Assessment Regime. In C. H. Tienken & C. A. Mullen (Eds.), *Education Policy Perils* (pp. 114–145). Abingdon: Routledge.

Sjöberg, S. (2019). The PISA-Syndrome – How the OECD Has Hijacked the Way We Perceive Pupils, Schools and Education. *Confero: Essays on Education, Philosophy and Politics, 7*(1), 12–65.