# Dire Straits-Education Reforms

## Ideology, Vested Interests and Evidence

Montserrat Gomendio
and Jose Ignacio Wert

https://www.openbookpublishers.com





ISBN Paperback: 978-1-80064-930-9

ISBN Hardback: 978-1-80064-931-6

ISBN Digital (PDF): 978-1-80064-932-3

ISBN Digital ebook (EPUB): 978-1-80064-933-0

ISBN Digital ebook (AZW3): 978-1-80064-934-7

ISBN XML: 978-1-80064-935-4

ISBN HTML: 978-1-80064-936-1

DOI: 10.11647/OBP.0332

Cover image: Kimberly Farmer, A collection of books (2017), https://unsplash.com/photos/lUaaKCUANVI. Cover design: Jeevanjot Kaur Nagpal.

# 7. Education Reforms

## The Interaction between Ideology, Governance, Conflicts of Interest and Evidence

### 7.1. Who Cares about Education and Why?

Education reforms matter. They actually matter a great deal, both at the individual and microsocial levels, as well as at the societal level. The reason for this is that good education systems equip individuals with the knowledge and skills required to obtain good-quality jobs, to navigate uncertain and rapidly changing labour markets, to develop innovative ways to deal with new challenges and to integrate fully in complex societies. Over time it has become increasingly clear that it is not just about having an education, but rather about enjoying a good-quality education, as knowledge-based societies and technological change have resulted in dramatic increases in the demand for higher levels of skills that are likely to continue or even accelerate in the future. Well-educated people can take control of their lives and adapt to a changing landscape, while poorly-educated people will struggle. At the aggregate level, high-quality education systems are key for improving human capital, which is the fundamental driver of economic growth and prosperity.

Families care deeply about education because it has a strong influence on the future of their children, but how much education is valued, and the extent to which parents get involved, varies significantly between countries and cultures. Employers also value education highly since it ensures that students acquire the increasingly complex levels of skills

 https://doi.org/10.11647/OBP.0332.07

that they demand; obviously they care more in societies where advanced economies already demand high levels of skills. Political parties seize the opportunity that educational ideologies provide to feed their electoral bases with educational flags, which seem to have considerable mobilising influence. Governments can exert a strong influence through education: they can boost human capital and economic growth; promote social mobility and equity; instil common values; facilitate integration and social cohesion, thus shaping the fundamentals of their societies, which could result in strengthening citizens' support. Finally, the sheer size and complexity of education systems and the levels of investment that they require imply that there are many stakeholders who receive direct benefits from the education system. The most numerous and politically powerful are teachers, who are expected to share the goal of achieving a high-quality education system in which they will thrive.

In this book, we try to understand the following paradox: everyone seems to agree on the value of high-quality education, and yet education reforms are uniquely difficult to implement. What are the main obstacles and how might they be overcome?

Conventional wisdom has it that governments rarely decide to embark on education reforms because the political costs in the short term are huge, while the benefits are rather diffuse and only become tangible in the long term. While it is obvious that it takes years before an education reform has an impact on all schoolchildren, and even longer before any benefits in adult life can be assessed, we still contest the argument that benefits only happen in the long term. As we have seen, many education reforms do have a positive impact in the short term. Thus, the main issue seems to be understanding the nature of the political costs and whether these can be overcome.

## 7.2. Ideological Battles: What Are They About?

In most countries, education is a highly polarised issue in the political arena. In this battlefield the quality of the education system as a whole plays a very minor role, partly because for most voters it seems to be a rather vague and wide-ranging concept which bears little relation to their specific needs, interests and values. The most divisive ideological flags focus on two issues.

On the one hand, parental choice is a key component for political parties on the right of the spectrum, and is attractive to parents who wish to have the freedom to choose the school they believe is best for their children in terms of quality, reputation, discipline, specialisation, values and in some cases religious faith. The education system can only allow parents to make meaningful choices if there is a diversity of schools on offer and this is achieved mainly through privately managed, government-funded schools (also known as charter schools). Thus, freedom of choice implies (i) the existence of different types of schools which, in turn, leads to competition to attract students, and (ii) transparency and accountability on the outcomes, so that parents can make informed choices.

On the other hand, equity is the main issue for political parties on the left of the spectrum. In this case, the focus is on public education as a guarantor that all children will have access to the same educational opportunities and the emphasis is on comprehensive policies which ensure that all children are treated equally.

Many education systems show that parental choice and equity are compatible, but ideological battles require much more than the defence of certain principles. The demonisation of the principles defended by one's opponents is just as important. In this rather destructive process, parental choice is portrayed by the left as leading to the privatisation and marketisation of education, promoting segregation between children based on their socio-economic status, taking resources away from the public system and thus overwhelming public schools with disadvantaged students who face greater challenges. In contrast, political parties on the right caricature the pledge for equity as a disguised attempt to conceal the limitations of a public system too weak to deal with any form of competition and too mediocre to generate anything but low standards for all students. As a consequence, the original priorities become blurred and the debate often ends in an oversimplistic and false dichotomy between private versus public education, which are portrayed as serving the interests of the elites or those of the most vulnerable parts of society respectively. These, of course, are strawmen. But they are powerful images that define the coordinates along which the educational debate unfolds in many places.

These basic ideological views have overspill effects that end up influencing political attitudes on many other educational issues. Raising equity to the status of the main, if not the only, goal of the education system, often goes hand-in-hand with supporting the abolition of any measures that are regarded as 'discriminatory' or leading to 'segregation', such as grouping students by their ability so that teachers can achieve greater learning gains with a more homogeneous group of students, or tracking of students into academic and vocational paths before they enter upper-secondary education. Thus, advocates of equity as the main goal of education systems tend to support the implementation of the most radical comprehensive policies, even though the evidence shows that, in societies with high levels of inequality, such education policies can lead to the worst outcomes in terms of equity. In other words, while egalitarian societies have implemented comprehensive policies successfully, in societies with high levels of inequity the levels of student heterogeneity are so high that they require other policies to deal with this major challenge first and foremost.

The trend to avoid any policy which could potentially be linked to some form of discrimination has become so influential in recent years that standardised student assessments are being questioned (or eradicated) because socio-economic status, immigrant status, gender and ethnicity tend to have an impact on the results. It is obvious that this is a symptom of the extent to which the education system has failed to minimise the impact of those factors, not a causal effect. Standardised assessments are just metrics which set common targets and the results reflect the extent to which the education system has enabled all students to achieve such goals. But many now seem to believe that being blind to this reality will make the problem disappear. Ideology is based on beliefs, fears and values, and its links to evidence are tenuous at best.

Those in favour of parental choice also tend to support competition between schools, accountability, transparency and a culture of effort. In other words, diversification and merit rather than uniformity. The risk in this case is that, unless clear rules are established, privately managed, government-funded schools may select students according to their socio-economic status, or discriminate against minorities, leading to social segregation. Furthermore, the main limitation of this ideological position is that, while it defends a parent's right to choose

what they regard as best, it does not make clear proposals about the overall education system, thus catering to a very targeted audience. In the worst-case scenario, it could lead to parents choosing a few good schools, while the public system as a whole remains in a poor state. Also, information asymmetry has to be considered in order to achieve a level playing field, since some parents have access to rich information resources on the quality of schools, while others do not.

While these issues seem to constitute the main pillars of the ideological views at opposite ends of the spectrum, in many countries the fiercest battles focus on issues which have less to do with the quality of the education system. In countries where no common narrative of history has prevailed, or where there are strong pro-independence parties in some regions, the curricular content of subjects such as history is the source of constant and bitter disputes. Those issues that are related to strong values and beliefs held by different groups will generate an ongoing conflict as to which will predominate or how they can be reconciled; in addition, as perceptions on culturally sensitive issues (such as abortion, euthanasia or LGBTQ+) evolve over time, the curriculum needs to adapt and the time lag tends to generate constant tensions. One clear example is religious faith: an agreement has to be reached as to whether different schools will adhere to different faiths, whether they will all teach a common subject that will deal with all religions and reflect their differences, or whether religion should remain entirely outside the scope of what schools teach.

Of course, ideological issues will always play a role in education reforms, in one way or another. They can be dealt with in different ways. In many instances, parties will be confrontational. They will claim that they will look for consensus but will do exactly the opposite, defending the radical views of those voters for whom education is a major electoral issue and trying to distance themselves from their opponents. This seems to be the only way in which political parties in most countries may obtain additional votes, by making education a major issue. In contrast, in countries where there is a *consociational* political culture, they will try find common ground and sideline the most divisive issues.

Ideological divides create deep cracks between political parties and between groups of voters that may be difficult to overcome after the elections are over. As a consequence, where those divides are deeper,

governments may only dare to embrace an education reform when they enjoy a parliamentary majority. However, as different parties alternate in power over time, a succession of reforms and reversals following different ideological interests is a common and sterile experience. Countries where an agreement can be reached between the main political parties over basic issues linked to improved student performance—and not influenced by ideological views—which will be implemented in the long run, are rare but meaningful exceptions.

The recent trend in the political landscape, which has seen many countries moving away from mostly bipartisan or quasi-bipartisan systems to a much more fragmented political landscape, does not seem to have made things easier for education reforms, since coalitions often avoid divisive political issues.

At the end of the day, what seems to be the most productive strategy—agreement on the basics and *agreeing to disagree* on the rest—is becoming more difficult as polarisation, identity politics and echo chambers are intensifying everywhere. Ideological confrontations on education are here to stay and will continue to obstruct performance-oriented education reforms.

## 7.3. Reform: Who Decides What, Who Funds, Who Has a Say, Who Has the Power to Block

When governments do decide to engage in an education reform, the governance arrangements in place determine who decides what, who raises and who spends the funding, who is accountable to whom, and the power that stakeholders and groups with vested interests have in supporting or blocking reforms.

Centralised and federal countries represent the two opposite extremes where there is clear definition of responsibilities at different levels: either central government or regions (respectively) hold most decision-making power, as well as the responsibilities for raising and allocating funding. The challenges for these two types of governance systems are very different. Centralised countries need to implement the right mechanisms to deal with the diversity in their countries. On the other hand, federal countries need to find ways to balance regional inequalities to prevent the education system from magnifying them;

this is normally achieved through additional funding provided by central government to compensate for regional disparities in wealth and resources, and by agreeing on common standards for all students to ensure that the education system implements the mechanisms required to avoid regional disparities in student outcomes.

In the last decades, there has been a widespread trend of decentralising education systems and the rationale has been that decision-making should take into account the specificity of regional and local needs, so that greater responsiveness to such needs would make the system more efficient. However, to achieve greater efficiency, several conditions need to be met, such as capacity building, a clear definition of responsibilities, and the implementation of accountability mechanisms, which unfortunately have not always been put in place.

Such decentralised, non-federal systems inhabit a grey area where responsibilities and funding can be shared in many different ways. Generally speaking, central government remains in charge of raising most of the funding through taxes which is then transferred to regional or local authorities, who decide how to spend it. If proper accountability mechanisms are not put in place, the disconnect between the responsibility for fundraising and the capacity to spend funds may lead to overspending and inefficient allocation of resources. It is also common for central government to retain decision-making powers on the general rules and architecture of the education system, as well as the national core curriculum and standardised national assessments, while the management of schools as well as some degree of autonomy in terms of curriculum and assessments, is transferred to regions,.

Among these *de facto* decentralised systems, a large number of actors play a role in education reforms and therefore the degree of complexity is much greater. In particular, central government has responsibility for approving new laws, but regions are key in the implementation of such reforms. Since it is often the case that different political parties hold power in central government and in different regions, political alliances may play a key role in the level of support or rejection that regions voice about specific reforms and, even more importantly, in whether they are willing to implement the reforms approved by the national parliament.

When responsibilities are not well-defined, there is an inevitable tension between regions demanding more resources and power and

central government reinforcing accountability mechanisms. Since most education reforms entail major re-distributional effects, regions are likely to react by escalating their demands. In this context, measures which may benefit the education system as a whole may be rejected by regions, either because they interpret them as a form of re-centralisation (e.g. national standardised exams), or because some regions regard themselves as losers in the new configuration (either because they fear receiving fewer resources, losing decision-making power, or being subject to greater control from central government). Although the interests of both central and regional governments are legitimate, central government's role is to improve the whole education system, while the role of regional governments is to ensure local advancement. Thus, the prism is very different and often leads to conflicts. In this context, voters' capacity to understand who is responsible for what, in the face of either improvements or declines, is often limited, particularly when the division of responsibilities is ill-defined.

As education systems have expanded to provide first universal access to education and then an increasing number of years of compulsory education, they have created complex and vast networks of schools which require a substantial amount of public funding. This growth has been accompanied by the surge of an increasing number of stakeholders who play very different roles, have different interests and hold very asymmetric powers. Obviously, policymakers would benefit from involving stakeholders in the design and implementation of education reforms since they can provide useful information, their support will give legitimacy to reforms, ensure their smooth implementation and increase the chances that they will last in the long term. Following this logic, it has become a mantra that reforms should be based on consensus. This seems an idealistic goal since major conflicts of interest often make consensus impossible, forcing governments either to adopt the lowest common denominator or to choose between the interests of powerful stakeholders and those of students. Thus, it seems more realistic to ask to what extent should different stakeholders be involved in the reform process and whether they should have veto powers.

Stakeholders can be categorised in two groups: those who benefit from a quality education system and those who benefit because they obtain resources directly. Clearly, the main beneficiaries of a quality

education system are the students. The purpose of education systems is to serve students and therefore their improved performance should be the *raison d'être* of every education reform. However, the influence of parents as stakeholders is very weak, mainly because they are not organised in ways that would make them powerful players, but also because they tend to lack information about the quality and efficiency of the education system as a whole. As stakeholders, parents hold the weakest position in the information chain, and very often they can be misled or confused about the consequences of policy interventions. Generally speaking, parents tend to mobilise when they perceive that their rights are being curtailed, i.e. when their freedom of choice is threatened either due to changes in admission rules, or because privately managed, government-funded schools are at risk. Employers also benefit from good education systems because this ensures that their workers have high levels of skills. However, in most countries employers only become directly involved in vocational education and training by providing on-the-job training, useful information about the needs of the labour market and, in some cases, by setting the standards of apprenticeships.

The education system mobilises a huge amount of public funding which provides resources to many providers, such as the textbook publishing industry, ed-tech companies, transport or school canteen companies. However, the majority of the funding by far is spent on staff (teacher salaries mainly). This means that teachers directly receive most of the funding that is allocated to education and their huge numbers imply that in many countries the public education payroll makes governments the largest employer. A natural consequence of this is that teachers have become organised in unions which defend their interests. During the era of institutional formation (when access to education expands and children spend an increasing number of years in school) there is strong alignment between the interests of governments and teacher unions: the funding goes to building and equipping schools, and mostly to hiring an increasing number of teachers.

However, once this stage is over most governments look for ways to improve the quality and efficiency of the education system. In many countries this involves decentralisation, choice, accountability and,

above all, a focus on student performance. This shift inevitably leads to conflicts of interest with unions.

Perhaps the main conflict arises as a consequence of the slowdown in the hiring of new teachers. The power of unions rests with the size of their membership and potential candidates (such as teachers on temporary contracts) expect unions to facilitate their entrance into the profession. Thus, once the stage where the education system is being built is over, unions shift their focus to class size: they demand smaller class sizes in order to ensure that more teachers continue to be hired. Although this is a legitimate defence of their interests, what seems misleading is that they disguise these vested interests as altruistic attempts to improve the quality of teaching.

Here, we have a very telling example of information asymmetry. Parents tend to believe that small class sizes equate to a more personalised, better-quality education, so they tend to actively support this demand. This has led to a widespread trend of gradually decreasing class sizes, despite the well-documented fact that this has no impact on student performance and is very costly. This has created a vicious cycle which has long-term consequences, since an increasing amount of funding goes to the salaries of a growing teaching force, instead of those resources being used to train and select fewer teachers of higher quality.

In many countries, unions have opted to ensure job security and good working conditions for all of their members, which entails a strong defence of similar salaries and job safety for all teachers (irrespective of their performance) and strong opposition to the dismissal of underperforming teachers, performance-linked pay, demanding training for teachers, selective hiring of teachers based on merit, and teacher evaluations. When unions have taken these demands to the extreme and they have enjoyed the political power to force governments to accept them, the result has been entrenched low teacher quality. This leads to a very dangerous loop since, given that the role of unions is to defend the interests of their members, when teacher quality is low unions will regard any accountability measure as a threat to their members. Thus, as governments have shifted the focus from inputs (resources) to outputs (student performance) and have looked for efficiency and accountability, conflicts of interest with the unions have

escalated. In those countries where unions have veto powers, they have systematically blocked reforms.

In the case of education reforms, the crux of the matter is that unions and other stakeholders are well-organised and have political power to defend the status quo, while those who would benefit from better outcomes (students and their families) are not.

For this reason, many governments face very difficult choices when they wish to undertake education reforms because they have to confront and deal with such intense conflicts of interest. Imposing an education reform is likely to shorten its life-span, but reaching a consensus often involves capitulating to vested interests and giving up on contentious but necessary policies which are rejected by ideological opponents. As we have seen, PISA claims that the policy recommendations that it provides in practice lower the costs for policymakers who are willing to follow them, because they can justify that reforms are based on evidence and are therefore exempt from ideological biases and free from manipulation by groups with vested interests.

Is the evidence from international surveys robust enough to empower policymakers to the extent that they can overcome conflicts of interest and ideological battles? Our answer is unfortunately not.

## 7.4. International Evidence versus Conflicts of Interest: Who is David and Who is Goliath?

International surveys have generated a vast amount of data on student performance in different countries which have revealed major differences in quality between countries and allowed comparative analyses to identify which features of education systems are associated with good outcomes. The availability of this information has raised hopes that governments could improve their education systems by using this objective evidence to design their education reforms and policies. In this way policymaking would ensure its success by using rigorous and irrefutable data, rather than partisan wrangling.

For this evidence-based approach to work, several conditions have to be met. The main condition is that the data must be reliable and that robust and solid conclusions can be drawn from it about 'what works' and in which contexts. The second is that all stakeholders (not just

policymakers) should agree that whenever there is strong evidence concerning the positive impact of specific policies, this should be enough to overcome underlying conflicts of interest, which should be sacrificed for the sake of improving student performance. Finally, for evidence to become a guiding light, societies should agree on what they wish to achieve through education.

All international surveys measure student outcomes in the same domains: reading, science and mathematics. While PIRLS and TIMSS also examine curricular content to analyse the extent to which students learn what they are taught in school, PISA obtains additional information through questionnaires and claims to measure the extent to which students are able to apply knowledge to solve problems in unfamiliar contexts, irrespective of whether the learning takes place in schools, homes or the social environment. Another relevant difference between these surveys is that while PIRLS and TIMSS assess student performance in specific grades, PISA evaluates the performance of fifteen-year-olds irrespective of grade.

Despite the more tenuous links between school practices and student performance in PISA, this is the only survey that has defined advising governments on good practices as its main priority and boasts about its influential role in education policy. Although the data provided by these surveys are not adequate to draw causal inferences, PISA has become a powerful tool in the political debate. Its influence arises mainly from the emphasis it places on targeting the media in most countries to enhance awareness of the national results and as an effective way to put pressure on governments to follow the policies that PISA recommends. Thus, in a way PISA has become part of the political debate rather than a source of independent evidence to allow governments to steer away from ideological battles.

In order to understand whether PISA policy recommendations are based on robust and solid evidence it is important to explain that they are based almost exclusively on correlations which are included in PISA publications and which follow a very similar pattern cycle after cycle, although the number of participating countries has grown over time. It is well-known that correlations do not allow the establishment of causal relationships, so this is a major weakness of PISA's conclusions. Although more sophisticated analyses have been carried out using the

available data from all ILSAs to establish more robust links between education policies and student outcomes, these are often ignored by PISA and instead remain within the academic realm, thus having no substantial impact on the media or policymakers. We therefore need to distinguish between the conclusions drawn from the correlations used by PISA and other, more robust analyses.

There is complete unanimity on one issue: the overall level of investment per student does not have any impact on student performance when it is above a certain threshold. While most OECD countries are above this threshold, many low- and middle-income countries are still below it. This result is robust when comparisons at different levels are made: comparative analyses between countries, as well as between regions within countries (which share the same architecture of the education system and often the same curriculum and assessments, thus eliminating many confounding factors that are present when countries are compared). In addition, when trends over time are analysed, it becomes clear that most countries have increased investment in education substantially over time, with no impact whatsoever on student outcomes.

The evidence that—above a certain threshold—investment *per se* is unrelated to student outcomes is the most solid evidence about what does not work in education. These findings contradict the most widely accepted premise in any debate on education: the higher the input (investment) the better the outcome (student performance). And the reverse: that budget cuts in education will inevitably lead to a decline in student outcomes. Knowing that the assumption is wrong does not seem to have had any impact on the education debate or investment policies. Why?

Since most of the investment in education is allocated to staff, the total amount is largely the product of two factors: the number of teachers and their salaries. In turn, the number of teachers is the result of the number of students and class size. The reason why the level of investment per student is unrelated to student performance is that neither class size nor teacher salaries have a direct impact on student outcomes. But the main driver is class size because it determines to a large extent the number of teachers and therefore has a huge impact on overall levels of investment.

It seems fair to say that this solid conclusion has had no impact at all as it goes against the tide of public conversation. Most countries have continued to substantially increase levels of investment over time, mainly because class size has continued to decrease, but also because teacher salaries have increased. There are several reasons why governments have been oblivious to this finding. The idea that investment is what matters the most in education is so ingrained that the majority of voters tend to agree irrespective of their ideology or political affinity and support growing levels of investment, so political parties in general do not dare question such a crucial matter. In virtually every country the political cost of reducing investment in education is huge, since it is interpreted as an unequivocal sign that a government does not regard education as a priority.

For unions and their allies on the left of the political spectrum, these two variables are the most important by far: decreases in class size require the hiring of more teachers and therefore make unions more powerful, and increases in teacher salaries greatly benefit their members. Thus, unions will go to great lengths to put pressure on governments and political parties to ensure that class size continues to decrease and teacher salaries continue to increase. Often, their support of reforms is contingent on increases in investment which ensure the continuation of these trends. The narrative built to defend these measures has been carefully crafted to avoid mentioning these vested interests and instead focuses on presumed benefits for students.

Despite all the evidence to the contrary, this narrative seems convincing to most families, owing to the widespread belief that smaller class sizes allow individualised teaching, which is assumed to improve quality of teaching. Policy recommendations from PISA have confused matters further: the conclusion in most cycles made it clear that neither investment, nor class size or teacher salaries, had a positive impact on student outcomes, but in the last cycle this stance was changed without any clear empirical evidence.

It seems as if PISA has decided that since the world was paying no attention to its most robust conclusion, it was better to accommodate 'mainstream beliefs' than to be seen as having no influence after two decades. This leaves governments with little support from PISA to defend one of the few policies actually based on strong evidence. This

has far-reaching consequences since, in the context of tight budgets, most investment continues to go towards reducing class sizes, rather than measures which would improve teacher quality, such as better teacher training and professional development, or incentives for teachers to perform well.

There is a second group of policy recommendations where PISA deviates from conclusions based on more robust analyses. As we have seen, there is evidence-based consensus that curriculum-based, standardised exit exams improve student performance. The reasons seem obvious: assessments in primary school allow early detection of those students lagging behind at a point when support measures are more likely to work, and assessments at the end of lower- and upper-secondary school are powerful indicators for students and teachers of what is expected and the level of effort required, giving ample opportunity to establish compensatory measures that allow disadvantaged students to achieve those standards, as well as providing guidance for further direction in students' educational pathways. Such assessments also ensure that teachers do not set different standards, and that education systems implement the necessary mechanisms to ensure that students in different regions achieve the same basic standards.

Although PISA did show in its early cycles that student assessments had a positive impact on student performance, it gradually changed its stance, aligning itself with those who believe that assessments are too stressful and could demotivate disadvantaged students who may feel that the standards are unachievable. As a consequence, it started warning against "high-stakes" exams (a negatively-loaded tag) and supporting assessments with no academic consequences, including those which only target a sample of students. At present, there is a tidal wave of rejection of student exams on the grounds that they are discriminatory, because factors such as socio-economic background, gender or migrant status tend to have an impact on the results.

In our view, student assessments are the equivalent of an X-ray which clearly diagnoses the weaknesses which afflict any education system but are not the cause of discrimination. Education systems will only be able to address such problems if they have a clear view, which can only be obtained if all students are assessed using the same metrics. The alternative would be to eliminate exams and therefore make education systems blind

to student outcomes, which would lead to governments renouncing their responsibility to minimise the impact of such factors. They would also give up on evaluating the impact of policies and on implementing any sort of accountability mechanisms based on student outcomes.

The repudiation of standardised and external student assessments is also the consequence of the stern defence of vested interests. In countries where teacher quality is poor, unions oppose student assessments because they fear that they will be used as indirect means of teacher evaluation. In such contexts, political parties on the left of the spectrum often decry exams using the argument that the fact that disadvantaged students or migrant students achieve poor results is proof that they have not enjoyed the same opportunities and therefore cannot be evaluated with the same metrics. In addition, in decentralised countries, regions often reject national assessments as a form of 're-centralisation', since they wish to expand their educational autonomy and minimise any form of accountability. In sum, there is a constellation of factors which make the evidence irrelevant despite its robustness.

Most analyses also show that giving more autonomy to schools in exchange for accountability has a positive impact on student outcomes. The reason is that when principals are able to make decisions about their schools, and teachers have the responsibility to choose which approach to use in the classroom, they tend to become more efficient thanks to their knowledge of their students and their needs. It also makes their jobs more stimulating, since they have greater responsibility and freedom to innovate. However, all analyses show that greater autonomy must go hand-in-hand with greater accountability, so that regional and national governments can make sure that the decisions taken by principals and teachers do lead to better student outcomes.

This is perhaps one of the more influential policy recommendations, although transferring greater autonomy to schools enjoys wider support than implementing accountability mechanisms, so the interplay between the two factors is often ignored. Granting more autonomy to schools has become part of PISA's recommended policy package, including for low-performing countries. This is unfortunate, since the evidence clearly shows that school autonomy will only bring benefits when principals and teachers are prepared to use those responsibilities in an effective way. This requires capacity building before responsibilities are transferred, as

well as a good-quality education system with highly-skilled principals and teachers. In fact, there is solid evidence that school autonomy has a negative impact on student outcomes in low- and middle-income countries.

Once again, unions fully support school autonomy as part of a more ambitious agenda that revolves around the idea that teachers should be 'trusted', a figure of speech which implies that teachers should have the freedom to choose what they teach and how, without being subject to accountability measures. The attitude of regions in decentralised systems is often contradictory, since they support greater autonomy as a means of further decentralisation, but when central government grants it, they often seize the opportunity to retain the new decision-making powers rather than transferring them to schools. Thus, while increasing school autonomy may be one of the clearest signals of the impact of evidence on education policy, it is often advised in contexts where it has harmful effects and it does not always go hand-in-hand with proper accountability measures.

This is a classic example of the risk of extrapolating practices which work in mature, high-quality education systems, and applying them in low-performing systems which are still not ready to take those steps. Improving education systems requires careful orchestration. School autonomy is one of the last steps in that sequence, because it relies on an established high quality of teaching, and on principals being true leaders. In other words, there is a mistaken logic which assumes that any practices present in top-performing systems will have a positive impact when transplanted into low-performing systems. The reality is much more complex. Only top-performing systems can successfully implement certain practices, such as school autonomy, because many other pieces of the complex puzzle are already in place for that change to have the desired positive impact. Thus, context matters, and most education policies cannot be extrapolated from high-performing systems to low-performing systems. But when vested interests benefit from such policies, this tilts the balance in favour of implementing them. It goes without saying that policy borrowing is not the same as policy learning.

The last policy recommendation on which PISA deviates from other analyses is the issue of school choice and the existence of privately managed, government-funded (charter) schools. A few countries have

traditionally developed these schools to allow parents to choose from a diverse offering. From the 1990s, a growing number of countries has introduced reforms to enhance school choice in order to make the education system more sensitive to the increasingly different needs of societies that have become more diverse and plural, and also to enhance quality and stimulate innovation. When addressing this issue, PISA tends to group privately managed, government-funded schools with private schools and directly compares this broad category with public schools. Thus, the analyses are not granular enough to compare the three categories individually and to draw clear conclusions about government-dependent private schools, which are the focus of much controversy. The general conclusion drawn by PISA is that private schools tend to do better, but this advantage disappears when socio-economic background is accounted for and therefore there are no clear benefits that could outweigh the risks associated with potential student segregation.

Thus, PISA seems more concerned about the potential risks of privately managed, government-funded schools leading to the segregation of students, than about any potential benefits. More robust analyses have concluded that privately managed, government-funded schools do generate better student outcomes due to their greater autonomy, which is linked to accountability of results, and to enhanced competition between schools to improve student performance in order to attract new parents. The data also show that government-dependent private schools tend to be much more cost-effective than public schools since they usually provide education at a lower cost per student than the latter. This is the result of a combination of factors: teachers in government-dependent private schools invest more time in teaching, these schools have larger class sizes, and principals have more control over the hiring of teachers.

However, when government-dependent private schools receive too little funding from government, they may not be able to afford to provide free education and instead charge tuition fees or add-on fees for extra-curricular activities. Since this undermines the principle of free-school choice, it is important that enough funding is provided by government and that these schools do not charge additional fees or follow a policy of selective admissions. Regulatory mechanisms should be implemented to prevent government-dependent private schools from targeting families

who can afford to pay for their children´s education and/or the best-performing students since both would increase inequalities. Thus, PISA warnings against the potential negative impact of school choice seem exaggerated, since the risks can and should be mitigated.

This policy measure is one of the most controversial in the political debate, because political parties on the right strongly support school choice while political parties on the left and unions strongly oppose it. The rationale used by political parties on the right is that school choice is a right that parents have and that privately managed, government-funded schools combine efficiency with quality. In contrast, unions and leftist parties claim that they segregate students according to socio-economic background and detract resources from the public system, which becomes over-burdened with disadvantaged students. Unions also fear that the large degree of autonomy granted to privately managed, government-funded schools and, in particular, the power that principals have to hire and dismiss teachers threaten some of their highly-valued privileges.

While in many countries this has become the most divisive issue, creating a schism between opponents of so-called 'privatisation' and supporters of choice and open competition, a few countries have managed to make it the core of a multi-party agreement. This is the case in the UK, where the main political parties have agreed to support a new type of school—academies—which entailed a major change in governance: the responsibility and funding shifted from local authorities to central government, and new accountability mechanisms based on student outcomes were put in place. Originally, low-performing public schools were converted into academies, but the model has proven so successful that it has grown, thanks to the support of consecutive governments of different ideological affiliations, to the extent that at present nearly 70% of publicly funded secondary schools are academies. It should be noted that agreement across political parties was possible, at least partly, due to the weakened power of unions in the UK.

The third and last group covers PISA policy recommendations not based on robust data or not based on data at all. Unfortunately, these policy recommendations address a major and fundamental dimension of education systems: equity.

There seems to be consensus on the idea that quality and equity are the two main dimensions of the education system. However, while measures of quality are quite straightforward (student performance), equity is multi-dimensional, and the available indicators only capture partial aspects. This leads to a largely unrecognised problem: when complex phenomena are simplified by the use of several indicators, this often gets lost in translation, meaning that one or a few of those narrow indicators is equated with the much broader and complex educational issue that they are meant to measure, i.e. equity. Thus, conclusions about such a multifaceted issue depend to a large extent on which indicator is chosen, and changes in one indicator are often used to draw general conclusions about progress towards educational equity. In short, equating any of these narrow indicators with the complex dimension of equity leads to the wrong conclusions.

PISA uses a large number of indicators to measure equity, such as the variance of student outcomes explained by student socio-economic background, the proportion of students who reach basic levels of proficiency, the variance explained by differences between schools and differences within schools, or the proportion of disadvantaged students that achieve high levels of performance (resilient students). As we will see, each indicator tells a different story and none of them tells the full story.

There is widespread consensus that socio-economic background is the factor with the single largest impact on student outcomes, a fact supported by robust analyses of data from different ILSAs. No education system has been able to completely overcome the influence of family background. While it seems unrealistic to expect that education systems will eliminate such an influential factor, it is important that they try to minimise it. However, good-quality education systems tend to raise the performance of all students, so disadvantaged students in high-quality education systems tend to outperform advantaged students in poor-quality education systems. From these findings, PISA concludes that the world is no longer divided between rich countries where all students perform better and poor countries where student performance is low. But unfortunately it is. Since poor students in Finland are not as poor as those in Colombia, if the wealth of the country is taken into account, it becomes clear that rich students in poor countries do perform worse

than poor students in rich countries. This is probably the consequence of systemic deficiencies affecting poorer countries and which parental resources cannot overcome, such as low curricular standards, teachers with low levels of skills and poorly designed assessments.

The broader and most challenging question is to what extent such differences between countries reflect the degree to which societies are egalitarian, or are mainly the result of the implementation of policies which simply minimise the impact of inequity. The evidence shows that the impact of household income upon student performance in countries with high levels of inequity is much greater than in more egalitarian societies. These findings suggest that education systems cannot overcome the impact of profound social and economic inequalities, unless policymakers succumb to the temptation of lowering standards for all students. They question the establishment of causal links between specific education policies and equitable outcomes, when these have been deployed in egalitarian societies. They warn against the risk of assuming that transferring policies which are implemented by egalitarian societies to countries with high levels of inequity will help to reduce inequality in student outcomes. The issue is clearly very complex, but the available evidence suggests that the degree of social and economic equity permeates education systems; as a result, egalitarian societies do not require major interventions against inequity. In contrast, societies with high levels of inequity face very different challenges and require specific policies to minimise the impact of inequality.

For most education systems, finding ways to deal with student heterogeneity is a major challenge, since teachers need to ensure that students with different performance levels continue to learn and achieve similar goals. Clearly, this issue is exacerbated in countries with high levels of inequality, where students have very different starting points, different levels of support at home and different access to resources. All these factors amplify differences in student performance, which become a major obstacle to learning gains. A number of policies have been devised to reduce variation in student ability when it compromises learning gains, but these have generated heated controversy. While supporters claim that teachers will be able to make greater learning progress in a classroom or group where students have similar levels of ability and needs, opponents argue that they will harm low-performing

students who will not be allowed to learn from their high-achieving peers, and that this will lead to discrimination and segregation since disadvantaged students will be placed in low-achieving groups/classes/tracks, irrespective of their level of performance.

The recommendations from PISA are consistent with this predominant narrative and do not support practices which aim to reduce student heterogeneity in performance, such as ability grouping, early tracking or grade repetition. However, careful analyses of the data reveal that such conclusions are not supported by robust evidence, so they must be challenged. In addition, the set of equity indicators that PISA uses is rather narrow and therefore fails to detect the harmful effects that supposedly 'egalitarian policies' inflict upon education systems in countries where social and economic inequity is pronounced. Vocational education and training, which PISA seems to abhor, is perhaps the clearest example of a policy conclusion contaminated by ideology and prejudice. One of the strongest policy recommendations from PISA is that VET decreases student performance and therefore should be delayed as much as possible. This is surprising, given that in most countries students cannot choose VET until the age of sixteen and PISA evaluates students at the age of fifteen. Put simply, PISA data cannot evaluate the impact of VET on student performance because there is no sample of students for the vast majority of countries.

In the case of grade repetition, PISA seems to fall into the well-known reverse causality trap: since the performance of students who repeat a grade is lower, then grade repetition lowers performance and should be avoided. This conclusion misses the point entirely because when students repeat a grade it is because their level of performance is so below average that they cannot continue to learn at the pace of their classmates. Grade repetition is a last resort and a second chance for students who have fallen far behind; thus, recommendations to abolish this practice do not address the root of the problem, which is the question of how to implement mechanisms earlier on that will allow students to catch up as soon as they start to struggle. Finally, conclusions regarding ability grouping face a similar issue: a simple correlation will show that this practice is more common when student performance levels are highly varied, as in non-egalitarian societies. The association

between these variables cannot be used as proof that ability grouping directly decreases levels of equity.

More importantly, PISA fails to grasp important issues which should be considered in this profoundly ideological debate on equity because it does not take into account any indicators that are not generated by the survey itself. The available evidence shows that practices which aim to reduce student heterogeneity and cater for different needs and interests, such as ability grouping and differentiated academic and VET programmes, do not decrease student performance. Furthermore, it shows that in non-egalitarian societies they tend to benefit low-performing students the most. When PISA recommendations are followed and 'comprehensive' policies borrowed from Nordic countries are implemented in countries with low-quality education systems and high levels of inequality, we see the worst outcomes in terms of equity: high rates of grade repetition, which lead to high rates of early school leaving. Students who drop out of school face high levels of unemployment for the rest of their lives. But PISA seems blind to these atrocious outcomes.

Obviously, we are not denying that any differential treatment of students carries a hidden risk of discrimination. Poorly designed ability grouping could result in students from low socio-economic backgrounds being unfairly assigned to low-performing groups, therefore limiting their chances of making progress. Similarly, old-fashioned VET systems may target students from underprivileged backgrounds and equip them with such a narrow set of skills that they can only aspire to low-skills jobs. The fear that education systems may fall into these traps is seemingly not supported by the evidence. But it is this fear that leads to recommendations to treat all students equally, which is widely regarded as an inclusive strategy. However, the needs of disadvantaged students will not be addressed by them receiving the same treatment as other students, because they require compensatory measures. While inclusive policies may serve as a safeguard against potential discrimination, they are by no means a solution to the very real problems that education systems face. When student heterogeneity becomes an obstacle to learning, reducing differences between students by grouping them according to their ability and offering different pathways grants the education system flexibility to adapt to the diverse needs of its student population.

## 7.5. The Geography of Education Success

A very consistent picture emerges when ILSA data are used to compare countries in terms of performance: the top performers are countries in East Asia, the low performers are mostly low- and middle-income countries in Latin America, Africa and the Near Middle East and the mid-performers are mostly European and North American countries along with New Zealand and Australia. The international surveys also reveal that in a number of cases differences between regions within countries are larger than differences between countries. Thus, despite their differences, ILSAs seem to be measuring similar features of student performance.

It is shocking that no strong narrative has been developed to explain the indisputable success of countries in East Asia, which are the real education superpowers. These countries have shown outstanding levels of performance from their earliest participation in ILSAs, and continue to improve over time. There is a commonly-held view based on the mistaken idea that students perform better only because of the extreme pressures that families exert on their children and the many hours of rote learning. This has led to the unfortunate conclusion that there is not much to be learned from these East Asian education systems in other parts of the world where education is not valued as highly and parents shy away from putting pressure on their children to perform well.

Instead, we argue that the fact that some of the policies implemented by these countries clearly contradict the dichotomous narrative (comprehensive vs segregation policies, trust in vs mistrust of teachers) that has invaded the ideological debate has downplayed any lessons to be learned.

Countries such as South Korea and Singapore clearly demonstrate that major improvements can take place much faster than is widely assumed, as long as the right policies are implemented and there is continuity over time. Continuity here does not mean preserving the same policies, but rather the opposite, i.e. that changes must take place over time as the quality of the education system improves, but these must be consistent with previous steps. Policymakers in these countries decided that human capital was the best asset, and in a few decades they evolved from illiterate societies to the top-performing education

systems in the world. One key element is the trade-off between teacher quality and class size. These countries, like most in East Asia, have opted for very large class sizes, so that investment goes mainly into teacher training and professional development. The excellent quality of the teaching force has yielded outstanding student outcomes. Such policy choices have been possible because unions do not have the power to block reforms, so admission processes both for university and for the teaching profession are highly selective. Teachers spend more hours engaged in professional development of excellent quality than in most other countries and this is linked to promotion and meaningful career choices. Because of the levels of excellence prevalent among teachers, the profession is highly valued and respected. But the misleading concept of trust does not apply: high curricular standards, as well as student assessments, are defined by central government and schools do not enjoy high levels of autonomy.

A key element of Singapore's sucess is the fact that tracking was initially implemented as early as primary level, in order to lower high rates of early school leaving. Given the success of these different tracks which cater to the needs of a diverse student population, Singapore has preserved tracking in lower-secondary education until today, long after it became a top performer. This fact has been conveniently ignored, probably because the success of early tracking contradicts the dominant narrative, which supports comprehensive policies.

In East Asia, consistency has been achieved over time because in some countries political parties adopt a very pragmatic view of education that avoids divisive ideological issues. More troubling, however, is the fact that other countries are semi-democracies (with restricted political competition) and some are outright authoritarian regimes. Recognising that limited political quarrelling and a lack of union veto powers on education policy seem to facilitate educational advancement does not mean that full democracies could not achieve the same results if ideological prejudices and vested interests were kept at bay.

The region which has moved in the opposite direction to East Asia is Latin America. Despite a better starting point around fifty years ago and huge efforts over the last decades to expand access to higher levels of educational attainment, the performance of students is very poor, so the returns in terms of skills are very low. This is mainly because

Latin American countries have implemented the opposite policies to those in East Asia. Investment has grown, but it has been directed mainly towards reducing class sizes, and few efforts to improve teacher quality have been successful. The power of the unions in Latin America is unparalleled, so they have put enormous pressures on governments to decrease class size (a measure that parents also support) and have forcefully rejected attempts to improve teacher quality by introducing selection measures for university degrees in education, training of a higher standard, and more demanding requirements for entering the teaching profession. As a consequence, these countries are locked in a downward spiral: unions defend the interests of a low-quality teaching force, such as lack of student assessments and teacher evaluations. Although class size continues to decrease, it has no impact on student outcomes.

In a region with huge levels of inequity, families, social movements and policymakers have high expectations of education's power as an engine of social mobility. This has led to the adoption of comprehensive policies and the abolition of so-called 'non-inclusive' policies. Thus, there is no ability grouping, no early tracking and VET is poorly developed. The universal aspiration is to access university, which is regarded as the only route to success. But these policies have failed in a big way. Comprehensive policies have not been able to deal with the huge diversity of students who enter school and have blocked any mechanisms which might deal with major differences in performance by providing alternative pathways. Consequently, high levels of grade repetition and early school leaving are prominent features of education systems in Latin America.

Most of the policy recommendations from PISA seem to originate from the contrast between Finland's unexpected success and the shock experienced by Germany in the very first PISA cycle (in 2000, two decades ago). The contrast between the unexpected success of a small and humble nation like Finland and the wounded pride of a powerful country such as Germany was crucial in the formation of a narrative which remains intact and very influential today, despite being—at least, in part—factually wrong. The Finnish education system was regarded as comprehensive (no early tracking), having a high-quality teaching force, high levels of school autonomy and achieving excellence by

prioritising equity. In contrast, Germany had early streaming of students into vocational and academic tracks from the age of ten years, and family socio-economic background and immigrant status had a much greater impact on student performance.

As with all influential narratives, this one is rather simplistic: education reforms should address inequalities before student excellence by designing comprehensive systems which do not segregate students into academic vs vocational tracks, ability grouping, or different types of schools (e.g., charter vs public). High levels of school autonomy and good-quality teachers led to what proved to be a far-reaching conclusion: teachers should be trusted. This concept was used by unions all over the world to argue that the evidence showed that teachers should not be evaluated, should not follow a pre-defined curriculum, and should have autonomy to innovate and decide what they teach and how they teach it. Some have taken a step further, arguing that teachers should own their profession and design education reforms, a stance supported by PISA.

But this narrative is not supported by evidence. While Finland became a top performer in reading among a small group of countries, in the following years the performance of its students declined. This suggests that some of the initial conclusions may be wrong. It also shows that success in PISA may impede necessary reforms when countries rely too much on their accomplishments. Alternative explanations for Finnish students' high performance in reading include policies implemented well before PISA 2000 in a much more centralised system where schools did not enjoy high levels of autonomy and teacher excellence was common. The time lag between the year 2000, when fifteen-year-olds were assessed, and the policies in place when they joined schools, is often ignored. In addition, families in Finland play a very important role in developing the reading skills of young children, so this may be an excellent example of PISA not acknowledging causal factors which are most visible outside schools.

In the small group of countries which experienced the so-called PISA-shock, such as Germany, evidence from successive cycles shows that, although in general they did implement reforms, none of these countries improved over time. This has far-reaching implications, since the strategy that PISA has systematically followed, which focuses on targeting the media to put pressure on governments to react, does

not seem to have worked, except in the case of PISA's own self-praise. Generating media outrage and placing governments between a rock and a hard place seems an odd strategy for a member-led organisation such as the OECD, where governments decide whether to join PISA and fund their countries' participation in addition to members' contributing to OECD core funding. It may also lead to unintended consequences, since it is possible that under such high pressure some governments may react too hastily and start reforms before they have the time to analyse the pros and cons of different options. Alternatively, since PISA claims to uncover the problems that an education system faces, then promotes a heated public debate and points fingers at governments, and finally claims to have the solutions to the very same issues that it has identified, some governments may end up following policy recommendations that do not apply to their specific contexts.

Those countries which improved after they started participating in ILSAs may provide more reliable insights as to which policies contributed to their success. These are all European countries which followed what seem to us the ABC of good practices: improve teacher quality, define a coherent curriculum with high standards, implement student assessments which are well-aligned with the curriculum, modernise and develop VET and, once a certain quality has been achieved, give more autonomy to schools in exchange for accountability. While ILSAs identified the positive trends in student performance over time in these reformist countries, it is deeply troubling that this evidence has not prevented most of these reforms from being reversed after they had proven to be successful. This fact shows that even when policies work, the evidence of this is not enough to protect them, because of a complete disregard for the objective assessments of policies.

## 7.6. Is the World a Better Place with Data?

According to PISA, no significant improvements in student performance have taken place almost two decades since the survey started. This is true when trends over time are considered for OECD countries, but also when a much larger group of countries is examined, since very few show improvements over time. This represents a failure of its self-proclaimed mission: to identify good practices, to advise governments

on which policies should be implemented and, in this way, to enhance student performance all over the world.

The OECD claims that PISA has helped policymakers lower the costs of implementing education reforms by backing difficult decisions with evidence, and therefore concludes that the lack of progress detected is the result of countries not implementing the right policies, by which it means the policies that the OECD recommends. But putting the blame on governments seems unfair and unsubstantiated.

We have argued in this book that it is more constructive to analyse the interaction between evidence, vested interests and ideology. The picture that emerges is much more complex. We identify three types of evidence depending on how robust the data are: strong, context-dependent, and weak. The first group includes variables related to levels of investment. The evidence from ILSAs is particularly strong regarding the lack of impact of overall investment, and its two main components (class size and teacher salaries), on student performance. But this evidence leads to a head-on clash with the vested interests of teacher unions, which benefit greatly from decreases in class size and increases in teacher salaries, which require higher levels of investment. Thus, when unions are powerful and the evidence generates conflicts with vested interests, the evidence does not play any role.

The second group includes policy recommendations which, based on the evidence provided by ILSAs, are strongly context-dependent: standardised student assessments have a positive influence if well-aligned with a high-quality curriculum; school choice does have a positive impact as long as it does not select students according to socio-economic status or demand fees from parents; school autonomy has a positive impact only among high-quality education systems and when implemented along with accountability mechanisms. Thus, it may be difficult for policymakers to evaluate which policies are required in their specific context.

Finally, the evidence concerning policy recommendations on equity is weak. This is partly because equity is multidimensional so all conclusions are partial and depend on which indicator is used. Furthermore, PISA does not take into account indicators which the survey itself does not generate and therefore misses crucial information by which to assess equity, such as rates of early school leaving. When

the evidence is weak, ideology takes over. It is worrying that this does not only happen at the level of political parties. It seems unfortunate that ideology also influences PISA's policy recommendations; in the absence of strong evidence, it seems to embrace the predominant narrative in a questionable effort to be seen as influential.

Despite these shortcomings, the evidence provided by ILSAs has proven very useful in identifying major differences in performance between countries and generated a wealth of data which are being used to decipher what seemed an intractable problem: how can education systems improve? These data represent the only way to understand which policies governments should implement to improve student outcomes in different contexts. But by targeting the media and causing such an uproar in the political debate, PISA has turned itself into a high-stakes exam for policymakers of the kind it no longer supports for students. As a consequence, governments expose themselves to huge media scrutiny, which may have a major influence on the way particular education policies or reforms are perceived by their societies. In exchange, PISA must ensure that its policy recommendations are based on strong evidence and that it is accountable when the reliability of the results generate reasonable doubts, if it wishes to be regarded as a trusted source of data.

Policymakers often face difficult decisions when confronted with a divisive ideological debate and powerful vested interests. It is often assumed that most choose not to act due to fear of political costs, but many may evaluate the situation and realistically conclude that entrenched conflicts of interest with powerful stakeholders make reform attempts unlikely to succeed. The only known fact is that those who do embrace education reforms and are willing to pay the political costs often encounter insurmountable obstacles. Since education systems serve students, progress will only be achieved when families and societies understand which polices benefit them, to the extent that civil society as a whole supports such changes. In the absence of a common understanding, major changes at the systemic level may prove impossible, and in this case the alternative may require taking small steps by implementing pilot studies which will eventually expand if proven to be successful. Small steps may lead to major changes, but it will take time. It remains an open question whether education systems can wait much longer for such change.