United Nations
Educational, Scientific and
Cultural Organization

UNESCO INSTITUTE FOR STATISTICS

SUSTAINABLE DEVELOPMENT GOALS

Handbook on Measuring Equity in Education

fhi360
THE SCIENCE OF IMPROVING LIVES

Oxford Policy Management

REAL
Research for Equitable Access and Learning

Handbook on Measuring Equity in Education

United Nations
Educational, Scientific and
Cultural Organization

UNESCO
INSTITUTE
FOR
STATISTICS

SUSTAINABLE
DEVELOPMENT
GOALS

UNESCO

The constitution of the United Nations Educational, Scientific and Cultural Organization (UNESCO) was adopted by 20 countries at the London Conference in November 1945 and entered into effect on 4 November 1946. The Organization currently has 195 Member States and 10 Associate Members.

The main objective of UNESCO is to contribute to peace and security in the world by promoting collaboration among nations through education, science, culture and communication in order to foster universal respect for justice, the rule of law, and the human rights and fundamental freedoms that are affirmed for the peoples of the world, without distinction of race, sex, language or religion, by the Charter of the United Nations.

To fulfil its mandate, UNESCO performs five principal functions: 1) prospective studies on education, science, culture and communication for tomorrow's world; 2) the advancement, transfer and sharing of knowledge through research, training and teaching activities; 3) standard-setting actions for the preparation and adoption of internal instruments and statutory recommendations; 4) expertise through technical cooperation to Member States for their development policies and projects; and 5) the exchange of specialized information.

UNESCO Institute for Statistics

The UNESCO Institute for Statistics (UIS) is the official source of cross-nationally-comparable data used to monitor progress towards the Sustainable Development Goal for education (SDG 4) and key targets related to science and culture. As confirmed in the Education 2030 Framework for Action, the UIS has the mandate to "work with partners to develop new indicators, statistical approaches and monitoring tools to better assess progress across targets related to UNESCO's mandate".

The UIS was established in 1999. It was created to improve UNESCO's statistical programme and to develop and deliver the timely, accurate and policy-relevant statistics needed in today's increasingly complex and rapidly changing social, political and economic environments.

Published in 2018 by:
UNESCO Institute for Statistics
P.O. Box 6128, Succursale Centre-Ville
Montreal, Quebec H3C 3J7
Canada
Tel: +1 514-343-6880
Email: uis.publications@unesco.org
http://www.uis.unesco.org

ISBN 978-92-9189-227-3
Ref: UIS/2018/ED/TD/1

© UNESCO-UIS 2018

This publication is available in Open Access under the Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) license (http://creativecommons.org/licenses/by-sa/3.0/igo/). By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository (https://en.unesco.org/open-access/terms-use-ccbysa-en).

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities or concerning the delimitation of its frontiers or boundaries.

The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.

Table of contents

List of abbreviations **7**

Foreword **9**

1. Introduction **11**
1.1 Relevance of equity in education 11
1.2 Knowledge gaps and intended use of the handbook 11
1.3 Overview of the handbook 13

2. Setting out a conceptual framework for measuring equity in learning **16**
2.1 Introduction 16
2.2 Equity in learning: A theoretical background 16
2.2.1 Why do we care about equity in learning? 16
2.2.2 What do we mean by equity? 17
2.3 Concepts for measuring equity 23
2.3.1 Meritocracy 23
2.3.2 Minimum standards 25
2.3.3 Impartiality 25
2.3.4 Equality of condition 30
2.3.5 Redistribution 32
2.4 Equality of what? 34
2.5 What should an equity measure look like? 38
2.5.1 Five desirable properties for equality of condition measures 38
2.5.2 Five desirable properties of impartiality measures 40
2.5.3 Desirable properties of redistributive and meritocratic measures 42
2.6 Summary 42

3. Proposed operationalisation of equity measurement **46**
3.1 Introduction 46
3.2 Visual representations of equality of condition 47
3.2.1 The histogram or the probability density function 47
3.2.2 The cumulative distribution 49
3.2.3 The Lorenz curve 51
3.2.4 Using visualisations of equality of condition 53
3.3 Measuring inequality: A catalogue of common metrics 53
3.3.1 Equality of condition 55
3.3.1.1 Differences 55
3.3.1.2 Ratios 58
3.3.1.3 Dispersion 58
3.3.1.4 Cumulative information 61

3.3.2 Impartiality 64
3.3.2.1 Differences 64
3.3.2.2 Ratios 66
3.3.2.3 Dispersion 66
3.3.2.4 Cumulative information 68
3.3.2.5 Analytic tools for testing impartiality 69
3.4 Availability and comparability of education data on equity dimensions 70
3.4.1 Scarcity of data on key dimensions of equity 71
3.4.2 Comparability of data on key dimensions of equity 73
3.5 Design and sampling considerations 74
3.6 Equity analysis: A proposed sequence 77

4. Measuring equity for national education planning 80
4.1 Measuring equity within national education plans 80
4.1.1 Methodology 81
4.1.2 Indicators included in national education plans 81
4.1.3 At what stages in the education cycle is equity in learning being measured? 88
4.2 Data needs for measuring equity 95
4.2.1 A more expansive approach to how disadvantage is measured 95
4.2.2 Expanding the coverage of data collection 101
4.2.3 Focusing explicitly on disadvantaged sub-groups from the earliest years 104
4.3 Conclusion 105

5. Redistribution of government financing to promote equity in education 108
5.1 Assessing who benefits from government spending on education 108
5.2 Formula funding strategies for redistribution of government education resources 113
5.2.1 Examples of formula funding 115
5.2.2 Formula funding needs to take account of both access and quality 118
5.2.3 Formula funding needs to take account of regional inequalities within decentralised systems 118
5.2.4 Redistributive formulae need to include teacher salaries 119
5.2.5 Schools need autonomy over the spending of resources, with guidance for using it in ways that address education quality for disadvantaged groups 119
5.2.6 Even where quality issues are addressed at school level, resources rarely reach disadvantaged groups 119
5.2.7 The effects of redistribution on narrowing learning gaps take time 120
5.3 Household spending on education: Implications for equity 121
5.4 National education accounts for tracking progress towards equitable financing 123
5.5 Conclusion 124

6. Concluding remarks 127

Annex A. Major international learning assessments 130

Annex B. 75 national education plans 134

List of figures

Figure 2.1 Imperfect and perfectly meritocratic distribution of university admissions 24
Figure 2.2 Proportion of children reaching three types of minimum standards 25
Figure 2.3 Impartiality of education based on test scores by sex, location and wealth quintile 28
Figure 2.4 Impartiality of education with respect to wealth 29
Figure 2.5 Lorenz curves showing a perfectly equal, and somewhat unequal, distribution of years of education 30
Figure 2.6 Federal government redistribution of education financing in Brazil 33
Figure 2.7 A model of how an education system works 35
Figure 3.1 Hypothetical test score distributions with varying levels of inequailty 48
Figure 3.2 Empirical distributions of PIRLS test scores for Canada Oman 49
Figure 3.3 Hypothetical cumulative distributions under different distributional assumptions 50
Figure 3.4 Empirical cumulative distributions of EGRA oral reading fluency (ORF) results for Haiti and Uganda 51
Figure 3.5 Hypothetical Lorenz curves illustrating two populations with different distributions of years of schooling 52
Figure 3.6 Lorenz curves for years of schooling in the United States and Burundi 53
Figure 3.7 Decision tree for equity analysis and applicable metrics 55
Figure 3.8 Hypothetical distributions of simulated test scores under different distributional assumptions 57
Figure 3.9 Gaussian distribution in relation to Chebychev's rule 59
Figure 3.10 PIRLS test scores from Morocco and Italy 61
Figure 3.11 Hypothetical Lorenz curve for years of schooling and Gini coefficient 62
Figure 3.12 Symmetric Lorenz curves for two independent populations with equal Gini coefficients 63
Figure 3.13 Between-country PIRLS achievement gaps 65
Figure 3.14 Adjusted gender parity index of the primary out-of-school rate, 2016 67
Figure 4.1 In education plans, participation and completion indicators feature more heavily than learning indicators 84
Figure 4.2 The majority of equity indicators for participation focus on sex 86
Figure 4.3 Among equity indicators on completion, sex dominates 87
Figure 4.4 Performance of 19 education plans against GPE quality standards 94
Figure 4.5 Educational opportunities vary greatly across rural India 97
Figure 4.6 Socioeconomic inequities are exacerbated by other disadvantages 100
Figure 4.7 In almost every country, poorer children are far less likely than richer children to complete primary school, latest year available (2006-2014) 102
Figure 4.8 Among poorer girls in rural Pakistan, those out of school are far less likely to be learning 103
Figure 4.9 Ensuring all 12-year-olds are learning the basics by 2030 will require efforts targeted at most disadvantaged 104
Figure 5.1 In most of 31 selected low- and lower-middle-income countries, the poorest children receive just a fraction of government expenditure on education 111

Figure 5.2 The poor-rich disparity in beneficiaries of government education expenditure is most extreme in tertiary education 112
Figure 5.3 Pro-rich biases in expenditure vary greatly across selected countries 114
Figure 5.4 Households fund a large share of education in some countries 122

List of tables

Table 2.1 Classification of equity concepts and related equity norms 23
Table 2.2 Selected impartiality measures 27
Table 2.3 Selected equality of condition measures 31
Table 2.4 Selected redistributive equity measures 34
Table 2.5 Levels of measurement 37
Table 3.1 Selected measures of inequality 54
Table 3.2 Common education indicators and their applicability to inequality measurement 56
Table 3.3 Availability of equity dimension data in international and regional databases and primary sources .. 72
Table 3.4 Expected domain size across three equity dimensions of different proportions 76
Table 4.1 Indicators included in national education plans 82
Table 4.2 Indicators of learning, by dimensions of equity, for lower grades of primary education in national education plans 90
Table 4.3 Indicators of learning, by dimensions of equity, for upper grades of primary education in national education plans 90
Table 4.4 Indicators of learning, by dimensions of equity, for secondary education in national education plans 91
Table 4.5 Frequency with which specific marginalised groups were cited in 19 education plans 93
Table 5.1 Selected middle-income country interventions to increase equity of education expenditure 116

List of boxes

Box 2.1 Defining learning 36
Box 4.1 The Global Partnerhip for Education's approach to measuring equity in education planning 93
Box 4.2 International initiatives in support of equity measurement 99
Box 5.1 Data needs for benefit incidence analysis 109

List of abbreviations

CDF	Cumulative distribution function
DGP	Data Generating Process
DHS	Demographic and Health Surveys
EFA	Education for All
EGRA	Early Grade Reading Assessment
EMIS	Education management information system
EPDC	Education Policy and Data Center
GE index	Generalised entropy index
GPE	Global Partnership for Education
GPI	Gender parity index
IDPs	Internally displaced persons
IIEP	UNESCO International Institute for Educational Planning
LLECE	Latin American Laboratory for Assessment of the Quality of Education
MDGs	Millennium Development Goals
NEA	National Education Account
NGO	Non-governmental organization
NHA	National Health Account
OECD	Organisation for Economic Co-operation and Development
OPM	Oxford Policy Management
ORF	Oral reading fluency
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PDF	Probability density function
PMF	Probability mass function
REAL	Research for Equitable Access and Learning
SDGs	Sustainable Development Goals
TIMSS	Trends in International Mathematics and Science Study
UIS	UNESCO Institute for Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNSD	United Nations Statistics Division
WIDE	World Inequality Database on Education

Foreword

We know that education is a fundamental human right. We know that without it, our lives – and indeed our world – would be greatly diminished. The collective progress that has been made over recent decades to get millions more students into the classroom is cause for celebration. But with so many challenges remaining – from concerns about whether they are actually learning to the educational exclusion of so many disadvantaged children – there is no room for complacency.

Now, as never before, we need to track progress on education in more detail. Where are the learning gaps? Who is still missing out on an education? And very importantly, why?

The Sustainable Development Goals (SDGs) provide the mandate for a strong focus on equity in education, aiming to ensure that the most disadvantaged children and young people have the same opportunities as others. SDG 4 demands an inclusive and equitable quality education for everyone, leaving no one behind. The challenge now is to provide the robust evidence, driven by solid data, which will enable the effective monitoring of progress on educational equity. With data currently available for less than one-half of the global indicators needed to track progress towards SDG 4, it is time to rise to that challenge.

This Handbook sets out, in practical terms, how this can be achieved. Produced by the UNESCO Institute for Statistics (UIS), in collaboration with FHI 360 Education Policy and Data Centre, Oxford Policy Management and the Research for Equitable Access and Learning (REAL) Centre at the University of Cambridge, it provides all those involved in the measurement of educational equity with not only the key conceptual frameworks but also the practical tools to do the job. With countries under pressure to deliver data on an unprecedented scale, the Handbook also recognises that no country can do this alone, making a strong case for greater cooperation and support across governments, donors and civil society.

The delivery of equitable quality education underpins the world's development goals, from poverty reduction to the promotion of peaceful and inclusive societies. We hope that this Handbook will help to translate the commitments made to equitable education into tangible action to monitor progress towards this crucial global ambition.

Silvia Montoya
Director, UNESCO Institute for Statistics

1. Introduction

BY CHIAO-LING CHIEN AND FRIEDRICH HUEBLER
UNESCO Institute for Statistics

1.1 RELEVANCE OF EQUITY IN EDUCATION

Education has long been recognised as a basic human right. It is a critically-important requisite for the productivity and well-being of individuals and for the economic and social development of entire societies. Because of this, the importance of equal access to education has been emphasised repeatedly in international conventions. The *Universal Declaration of Human Rights* of 1948 and the *International Covenant on Economic, Social and Cultural Rights* of 1966 state that education shall be equally accessible to all on the basis of merit and individual capability (United Nations, 2017b, 2017a). Access to education and learning outcomes should not be affected by circumstances outside of the control of individuals, such as gender, birthplace, ethnicity, religion, language, income, wealth or disability.

Beyond the issues of fairness and basic human rights, there is ample evidence demonstrating the economic and social benefits of education (UNESCO, 2014a). Work linked to the human capital theory and returns on investment in education has shown that increased educational attainment is associated with higher personal earnings, reduced poverty and higher growth rates of national income (Becker, 1975, 2002). Other studies have examined not only the economic but also the social benefits of education (McMahon, 2009; Stacey, 1998; and Vila, 2000). Having more years of education is associated with better health, reduced maternal and child mortality, fewer disaster-related deaths, less conflict and increased civic engagement, among other benefits.

With the adoption of the Sustainable Development Goals (SDGs) and the Education 2030 Framework for Action in 2015, equity has been placed at the heart of the international development agenda for the first time. In the domain of education, SDG 4 calls on all UN Member States to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all" (United Nations, 2015).

Several targets under SDG 4 aim for equal outcomes for all population groups, including girls and boys, and women and men, but also other groups. Gender parity was already a prominent target in the Millennium Development Goals (MDGs) adopted in 2000, but the SDGs go beyond this narrow focus. Target 4.5 is most explicit in its focus on equity and its determination to "eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations" (United Nations, 2015). Target 4.5 commits all UN Member States to addressing all forms of exclusion and inequalities in access, participation and learning outcomes, from early childhood to old age.

1.2 KNOWLEDGE GAPS AND INTENDED USE OF THE HANDBOOK

Greater equity and inclusion in education cannot be achieved without increased efforts to collect and analyse data on the most excluded segments of the population (UNESCO, 2014). Yet, three years after the adoption of the SDGs in 2015, education data are often still incomplete and many of the most

marginalised groups remain invisible in statistics at national and global levels.

Both administrative data systems and household surveys tend to lack data on certain populations. These include, for example, persons displaced by conflict, children in child labour and other vulnerable situations, nomadic populations or students attending non-standard forms of education. In addition, students with disabilities or with limited proficiency in the language of the assessment are being excluded from participation in cross-country learning assessments, including the Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), and Programme for International Student Assessment (PISA). Moreover, schools that are located in remote regions might also be excluded from those assessments (OECD, 2016; Schuelka, 2013). In household survey data, high variance in indicator estimates for small population groups – such as members of ethnic, linguistic and religious minorities – is another important challenge *(*UIS, 2016*)*. This lack of comprehensive data makes it difficult to identify groups that may not be reaping the full benefits from education because of restricted access and insufficient learning.

In addition, the indicator framework for SDG 4 has not been fully developed. Lessons learned from the experience of monitoring the Education for All (EFA) targets suggest that to track progress each SDG target should be measurable, with the associated indicators and data sources identifiable at an early stage (Rose, 2015). The Education 2030 Framework for Action therefore mandates the UNESCO Institute for Statistics (UIS) to work with partner organizations and experts on the development of new indicators, statistical approaches and monitoring tools for the assessment of progress towards SDG 4 (UNESCO, 2016).

This handbook, produced by the UIS in collaboration with FHI 360 Education Policy and Data Center, Oxford Policy Management, and the Research for Equitable Access and Learning (REAL) Centre at the University of Cambridge, is intended as a reference for analysis and interpretation of education data. It is aimed at professionals involved in the measurement and monitoring of equity in education, which includes not only those working on the SDGs but also any stakeholders in the field of education: technical staff in ministries of education and national statistical offices, education practitioners, members of non-governmental organizations (NGOs) active in the field of education and researchers. Although users of the handbook are expected to have basic statistical knowledge and some familiarity with equity issues and indicator calculations, the subsequent chapters develop and review basic material for the respective topics.

The handbook is inspired by the SDGs and Education 2030 but is not limited to an examination of the proposed indicator framework for the 2030 goals. Instead, it is designed to be suitable for any national analysis and monitoring of equity in education and progress towards national goals. The handbook is primarily concerned with national policymaking and focuses on inequalities within countries. While these inequalities must be addressed to achieve the education SDG, their elimination is a goal worth pursuing independent of the international development agenda.

There is already a large volume of work dedicated to measuring equity, and much of this work is founded in analysis of economic inequality (e.g. Atkinson, 1970; Atkinson and Marlier, 2010; Cowell, 2011; Dalton, 1920; and Roemer and Trannoy, 2016). Important milestones include the development of the Lorenz curve and the Gini coefficient more than 100 years ago. These and other approaches first used in economics were subsequently applied to health and education.

The issue of inequity in education has been examined from different angles, including inputs, processes, outputs and outcomes, as well as in various contexts (e.g. education systems, providers and learners). These issues have been covered in many of UNESCO's *Global Education Monitoring Reports* since 2002. The 2005 Report (UNESCO, 2004) included a framework for assessing education quality, with attention to equity, and the 2013/4 Report (UNESCO, 2014) provided an in-depth assessment of the ways in which teaching and learning processes need to change to leave no one behind. Other UNESCO publications have highlighted gender inequality (UNESCO, 2012; UIS, 2017) and examined the reasons for exclusion from education (UIS and UNICEF, 2015).

Data collection must be improved to allow identification of excluded groups and more precise calculation of indicators that can serve as evidence for the design of targeted policy interventions. Detailed advice on collection of data is beyond the scope of the present handbook, but it does make the point that high-quality data fit for disaggregation are an essential prerequisite for analysing equity.

Learning is a lifelong process and the measurement of equity in education should consider all ages and levels of education. Because inequity in education can accumulate over time, measurement must start in the earliest grades of a country's education system and even in pre-primary education. Indeed, the aspirations of SDG 4 are holistic and cover learning opportunities throughout the lifecycle, from early childhood to adulthood and old age. Focusing inequality research on a single level of education ignores the process of accumulating disadvantage throughout the education cycle. Therefore, it is necessary for education planners to take an integrated approach to investigating inequity accumulated at each transition point between education levels and to develop aligned policies and measures (Chien, Montjourides and van der Pol, 2016; Reisberg and Watson, 2010).

Even if the analysis focuses on existing disparities within a country, there is a need to define common metrics and standards to ensure reliability and international comparability of the results. Efforts to develop international standards in support of global monitoring are among the core elements of work in the context of the SDGs, but they also pose some of the biggest challenges (UIS, 2016). This handbook aims to contribute to the debate by proposing standard approaches and tools that could be used by all practitioners in the field.

1.3 OVERVIEW OF THE HANDBOOK

This handbook addresses some of the knowledge gaps outlined above. Specifically, it provides a conceptual framework for measuring equality in learning; offers methodological guidance on how to calculate and interpret indicators; and investigates the extent to which measuring equity in learning has been integrated into country policies, national planning and data collection and analysis.

Chapter 2 of the handbook presents a conceptual framework for equity analysis, with an emphasis on equity in learning. It begins with a summary of the philosophical literature on equity and highlights several related principles, including equality of opportunity and considerations of fairness and justice, as these relate to the distribution of education resources to compensate for unequal starting points. The chapter then proposes five categories for the classification of measures of equity: meritocracy, minimum standards, impartiality, equality of condition and redistribution. The chapter concludes with a summary of the desirable properties of equity measures.

Building on the conceptual underpinning presented in Chapter 2, Chapter 3 describes different methods for measuring equity in education. It focuses on key univariate and multivariate metrics and their respective advantages and disadvantages for two of the five categories described in Chapter 2: equality of condition and impartiality. Chapter 3 begins with

an overview of visual representations of equality of condition that can be used to gauge the degree of inequality in a dataset, among them histograms, probability density functions and the Lorenz curve. The chapter goes on to describe common metrics for measurement of inequality, organized by the kind of data to be analysed, the desired type of analysis and the type of equity measure. Chapter 3 concludes with an overview of data that can be used for analysis of equity, as well as some of the challenges that may be encountered along the way.

Chapter 4 moves away from the theoretical discussions in Chapters 2 and 3 and examines the role of equity measurement in 75 national education systems, in order to offer guidance to both policymakers and other stakeholders tasked with improving equity in education. The chapter begins with an analysis of national education plans from all geographic regions to identify the presence – or absence – of equity dimensions in indicators for monitoring of progress towards increased access and learning. Based on the findings, the chapter offers a series of recommendations for expanded data collection, with an increased focus on the identification of disadvantaged groups.

Chapter 5 discusses government spending as a means to increase equity in education. The chapter examines national data to assess which groups of the population benefit most from government education expenditure and describes formula funding as a way to redistribute resources to those with the greatest need. In this context, the role of household spending on education and the potential of national education accounts as a tool to identify and address inequities are also discussed.

Chapter 6 concludes the handbook with a summary of the main findings and recommendations for future work on national and international education statistics.

REFERENCES

Atkinson, A. B. (1970). "On the Measurement of Inequality". *Journal of Economic Theory.* Vol. 2, No. 3, pp. 244-263.

Atkinson, A. B. and E. Marlier (2010). *Analysing and Measuring Social Inclusion in a Global Context*. New York: United Nations Department of Economic and Social Affairs.

Becker, G. S. (1975). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research. http://www.nber.org/chapters/c3730.pdf

—— (2002). "The Age of Human Capital". In E. P. Lazear (ed.), *Education in the Twenty-First Century*. Stanford: Hoover Institution Press. Retrieved from http://media.hoover.org/sites/default/files/documents/0817928928_3.pdf

Chien, C.-L., P. Montjourides and H. van der Pol (2016). "Global Trends in Access to Post-Secondary Education". In Anna Mountford-Zimdars (ed.), *Access to Higher Education: Theoretical Perspectives and Contemporary Challenges*. Oxford: Taylor & Francis/Routledge.

Cowell, Frank (2011). *Measuring Inequality*. Oxford: Oxford University Press.

Dalton, Hugh (1920). "The Measurement of the Inequality of Incomes". *Economic Journal.* Vol. 30, No. 119, pp. 348-361.

McMahon, W. W. (2009). *Higher Learning, Greater Good: The Private and Social Benefits of Higher Education*. Baltimore: Johns Hopkins University Press.

Organisation for Economic Co-operation and Development (OECD) (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris: OECD.

Reisberg, L., and D. Watson (2010). “Access and Equity”. Chapter 11 in *Leadership for World-Class Universities. Challenges for Developing Countries*. Chestnut Hill: Boston College. http://www.gr.unicamp.br/ceav/revista/content/pdf/Watson_Reisberg-Access_and_Equity_en.pdf

Roemer, J. E. and A. Trannoy (2016). “Equality of Opportunity: Theory and Measurement”. *Journal of Economic Literature.* Vol. 54, No. 4, pp. 1288–1332. https://doi.org/10.1257/jel.20151206

Rose, P. (2015). “Three Lessons for Educational Quality in Post-2015 Goals and Targets: Clarity, Measurability and Equity”. *International Journal of Educational Development*. Vol. 40, pp. 289-296.

Schuelka, M. J. (2013) “Excluding Students with Disabilities from the Culture of Achievement: The Case of the TIMSS, PIRLS, and PISA”. *Journal of Education Policy.* Vol. 28, No. 2, pp. 216-230. https://doi.org/10.1080/02680939.2012.708789

Stacey, N. (1998). “Social Benefits of Education”. *The Annals of the American Academy of Political and Social Science*. Vol. 559, pp. 54-63.

UNESCO Institute for Statistics (UIS) (2016). *Sustainable Development Data Digest: Laying the Foundation to Measure Sustainable Development Goal 4*. Montreal: UIS. http://dx.doi.org/10.15220/978-92-9189-197-9-en

—— (2017). *eAtlas of Gender Inequality in Education*. Retrieved from http://www.tellmaps.com/uis/gender/

UNESCO Institute for Statistics (UIS) and United Nations Children’s Fund (UNICEF) (2015). *Fixing the Broken Promise of Education for All: Findings from the Global Initiative on Out-of-School Children*. Montreal: UIS. http://unesdoc.unesco.org/images/0023/002315/231511e.pdf

United Nations (2015). “Transforming Our World: The 2030 Agenda for Sustainable Development.” United Nations General Assembly resolution A/RES/70/1 (25 September). http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1

—— (2017a). “International Covenant on Economic, Social and Cultural Rights”. United Nations High Commissioner for Human Rights (OHCHR). Accessed July 31. http://www.ohchr.org/EN/ProfessionalInterest/Pages/CESCR.aspx

—— (2017b). “Universal Declaration of Human Rights”. United Nations. Accessed July 31. http://www.un.org/en/universal-declaration-human-rights/index.html

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2004). *Education for All: The Quality Imperative - EFA Global Monitoring Report 2005*. Paris: UNESCO. http://unesdoc.unesco.org/images/0013/001373/137333e.pdf

—— (2012). *World Atlas of Gender Equality in Education*. Paris: UNESCO. http://unesdoc.unesco.org/images/0021/002155/215522E.pdf

—— (2014). *Sustainable Development Begins with Education: How Education can Contribute to the Proposed Post-2015 Goals*. Paris: UNESCO. http://unesdoc.unesco.org/images/0023/002305/230508e.pdf

—— (2016). *Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4*. ED-2016/WS/28. Paris: UNESCO. http://unesdoc.unesco.org/Ulis/cgi-bin/ulis.pl?catno=245656

Vila, L. E. (2000). “The non-Monetary Benefits of Education”. *European Journal of Education.* Vol. 35, No. 1, pp. 21-32.

2. Setting out a conceptual framework for measuring equity in learning

BY STUART CAMERON, RACHITA DAGA AND RACHEL OUTHRED
Oxford Policy Management

2.1 INTRODUCTION

What does equity in learning mean? The long history of debate on the nature of equity and inequality in political philosophy and ethics suggests that there may be no single universally-convincing answer to this question. Equity is a political issue, and differences in political views will influence the aspects of equity in which we are interested. Thus, any effort to measure equity cannot be divorced from a normative framework about fairness and justice. This chapter aims to provide such a framework for analysis, focusing on principles that can attract broad agreement. We take several accepted approaches to understanding equity in learning and examine the implications of each approach for measuring equity.

These fundamental questions in measuring equity in learning are intended to provide a conceptual framework for the handbook. The chapter starts with broad principles and a brief survey of the philosophical literature on concepts of equity, considering how they apply to education in particular. Measures of educational equity can be classified into five categories: meritocracy, minimum standards, impartiality, equality of condition and redistribution. The chapter describes how each of these categories relates to concepts such as equality of opportunity in the philosophical debate. The meaning of these concepts in practice depends on whether we are looking at educational inputs, processes or outcomes, and it is therefore useful to present a simple classification of the education indicators used for equity analysis. There are a number of desirable properties that equity indicators can have. Considering whether a particular indicator fulfils these criteria helps us to decide whether to use it or not. The chapter ends by setting out some of these desirable properties and explaining briefly why they are useful.

2.2 EQUITY IN LEARNING: A THEORETICAL BACKGROUND

2.2.1 Why do we care about equity in learning?

It is increasingly recognised that learning levels within many countries are highly unequal. This may happen, for example, because of institutional features of school systems, such as early streaming, regional diversity in expenditure or political engagement, unequal access to education and drop-out rates, or unequal access to different types of provider (OECD, 2012). Among OECD countries, those with more equal learning outcomes also have better average learning outcomes, suggesting that appropriate interventions in the education sector may have positive effects on both equality and the quality of education (Pfeffer, 2015). Economic inequality is associated with the distribution of numeracy skills among adults, although the direction of causation is unclear (Van Damme, 2014).

Arguably, no society will ever reach total equality in the learning outcomes of every individual. Differences in learning outcomes may depend on individual differences in ability and motivation, as well as the type of background one comes from and the type of resources one has access to.

So, when should we become concerned about unequal learning outcomes? The answer to this is partly empirical, depending on negative consequences of different forms of inequality. However, it is also partly philosophical, depending on what forms of distribution of an important good we consider acceptable or justifiable. For both parts of the answer, we need to be able to characterise distributions of learning and of the inputs and resources that determine learning, in a nuanced way. We need to understand both what aspects of the distribution are empirically associated with positive or negative consequences for society and the economy, and what types of distribution can be characterised as unfair in political debate.

2.2.2 What do we mean by equity?

This handbook focuses on equity in learning. Equity and equality are contested terms, used differently by different people. Following Jacob and Holsinger (2008, p. 4) we define **equality** as "the state of being equal in terms of quantity, rank, status, value or degree", while **equity** "considers the social justice ramifications of education in relation to the fairness, justness and impartiality of its distribution at all levels or educational sub-sectors". We take equity to mean that *a distribution* is fair or justified. Equity involves a normative judgement of a distribution, but how people make that judgement will vary.

Both concepts can be operationalised in a wide variety of ways. Equality can be applied across individuals, groups or countries, and to different indicators. Equity can be applied with different theories of justice in mind and with different understandings of the wider ramifications of the distribution of education.

An emphasis on equity suggests that a particular distribution needs to be justified, with some combination of reference to abstract principles and concrete evidence. In this section, we present some of the principles on the nature of equity and fairness drawn from political philosophy and ethics literature. We will note political differences in which principles are seen as most compelling but also areas of broad agreement. For example, many people are likely to agree that equal access to primary education is important, while fewer would agree that higher education outcomes should be more equal in a particular context. We highlight the principles and frameworks that are likely to generate broad agreement and can therefore be recommended as most useful for measuring equity in education.

Equality of opportunity

A common approach to dividing up inequalities into those that can be justified and those that cannot is by applying the principle of equal opportunity. Equality of opportunity means that everyone should have the same opportunity to thrive, regardless of variations in the circumstances into which they are born. Having been granted such opportunities, however, their outcomes will still depend on how much effort they put in. Individuals are responsible for, and have control over, their effort, and so the portion of inequality in outcomes that arises from differences in effort is fair, while the portion that arises from differences in gender or parents' wealth is not fair. For "effort" we could substitute "ability", "intelligence", "propensity for hard work" and so on, depending on what characteristics we see as a fair basis for outcomes to vary.

Equality of opportunity is often posited as a more reasonable alternative to the idea of eradicating inequalities in outcomes altogether. Focusing only on inequality in outcomes is sometimes seen as denying the importance of individual responsibility and

choice, and overlooking the diversity of preferences and tastes (Phillips, 2004). In education, it may be unrealistic, for example, to expect all children to attain equal learning outcomes by the end of primary school. No matter how attentive the education system is to the needs of different learners, differences are likely to arise due to their pre-school experiences, abilities and personalities. There might also be a social cost associated with making everyone equal. For example, it might mean a less efficient economy or an education system less able to focus resources on the most able students.

Equal opportunity has become widely entrenched in national law and international rights instruments. It is at the "heart of many international human rights provisions, starting with the 1948 Universal Declaration of Human Rights. The 1989 Convention on the Rights of the Child establishes a binding obligation on governments to work towards fulfilling the right to education 'progressively and on the basis of equal opportunity' (United Nations, 1989, Article 28). The right to equal opportunity for education is also enshrined in most countries' national laws and constitutions" (UNESCO, 2010, pp. 135-6).

Equality of opportunity in education also lends itself well to empirical analysis. Roemer (1998) proposes an influential formulation for thinking about it in a measurable way: if we identify inequalities in access to education, and these inequalities can be traced back to differences in circumstances, such as one's parents' wealth, then we deduce that people have not had equal opportunities.[1] Roemer (2002) considers models with individuals who belong to different "types" (say, rich and poor) and proposes "as a simple measure of the morally relevant degree of effort, the quantile of the effort distribution for his type at which an individual sits" (Roemer, 2002, p. 458). The intention is to control for ways in which belonging to a different type would likely influence the effort one is willing to exert and treat these differences as morally arbitrary. This approach lends itself readily to existing statistical techniques, such as ordinary least-squares regression, and has been taken up by several education researchers examining inequalities in learning outcomes or attainment (e.g. Gamboa and Waltenberg, 2015; Ferreira and Gignoux, 2011).

Although the relationship between educational outcomes and the background of students is inherently interesting, we question whether this really measures equality of opportunity in education. Roemer's recommendation that we use an individual's position in the distribution of (some measure of) effort for his or her "type" as an indicator of his or her morally-relevant degree of motivation or industriousness is not innocuous. It assumes that all differences in distribution of effort between types are morally arbitrary, while individuals can be held responsible for their degree of effort relative to other members of their type. Both parts of this assumption can be questioned. In education, we are often concerned with children. It is difficult to justify morally the idea that a person's life chances should depend heavily on how industrious they were as children, before they had even reached an age of legal responsibility.[2] Finally, innate ability should be seen as a circumstance not of the child's own choosing, yet in practice the analysis of education using Roemer's equality of opportunity concept has not controlled for innate ability in any way (for example, through testing cognitive skills in early childhood). There is, thus, a disconnect between the moral philosophy of equality of opportunity and the attempts to measure it in education to date.

As the discussion of Roemer and other works demonstrates, the definition of equal opportunities and what constitutes these circumstances is

1 Note that this concept does not distinguish social and economic barriers from legal ones and so is more in line with Rawls' "fair equality of opportunity" than more restricted legalistic definitions that are also found in the political literature.

2 Indeed, Roemer and Trannoy (2016) seem to agree with this, expressing the view that "all inequality regarding children should be counted as due to circumstances and none to effort" (p. 1308).

contested. At one extreme, it can be defined merely as the absence of overt discrimination based on, for instance, gender or race. At the other extreme, the bad luck of being born with less talent or intelligence can be seen as a circumstance that unfairly affects opportunities. In other words, saying that we are interested in measuring equality of educational opportunities says little about what we intend to measure until we specify further which circumstances or characteristics are seen as unfair sources of differences in outcome.

Most proponents of equality of opportunity do not suggest that it is the *only* relevant criterion for judging the fair or right distribution of goods. In the equal opportunity theory of Roemer and Trannoy (2016), inequalities can never be fair if they are not due to effort, but some inequalities are not fair even though they *are* due to effort. Cohen (2009, cited in Roemer and Trannoy, 2016) suggests we should consider the strain that large outcome inequalities could place on social unity, which might mean reducing them further than demanded by an equal opportunity theory.

Even when taking a specific and well-defined conception of equality of opportunity, it may not be obvious how it should be operationalised when it comes to measuring change. For example, does it mean that schools should have equal inputs (per student) or that inputs should be allowed to vary to compensate for disadvantage of some communities? Should it apply to all levels of the education system? Or is the point that education systems should be structured in a way that ensures people have equal opportunities in work and life after they have left full-time education? These are questions to be addressed in national policy and international agreements.

Equality of opportunity is a central idea in inequality debates but needs to be specified carefully before it can be applied to measurement. This handbook avoids using this term to describe educational inequality indicators, because how one conceives of equality of opportunity is likely to be contested and also because there are likely to be several potential indicators that could be used to measure a particular conception of equality of opportunity.

Justice as fairness

Perhaps the most famous attempt at a more comprehensive definition of what types of inequality can be justified is the idea of justice as fairness described in John Rawls' book, *A Theory of Justice*. Rawls' theory of justice is based on the ideas of "society as a fair system of cooperation" and "citizens as free and equal persons" (Rawls, 1971). Reflecting these ideas, Rawls brings in the "veil of ignorance" as a device for thinking about the ideal society. We should think about what type of society we would want to be born into as if under a veil of ignorance, that is, as if we didn't know what type of person we would be born as, whether to rich or poor parents, intelligent or not, in a deprived rural area or a rich city.

Rawls argues that, if put into this hypothetical situation, rational actors would choose a society where inequalities would be accepted by the worst-off in society. The veil of ignorance would compel us to start from a presumption of total equality, reflecting the fundamental equality of citizens. Not knowing where we would end up in the social hierarchy, we would want to ensure that we had access to a set of basic liberties, such as freedom of person and the freedom from arbitrary arrests and seizures. When it comes to social and economic inequalities, we might permit some deviation from the starting point of total equality. But, Rawls argues, we would only do so under two conditions:

> First, they [the inequalities] are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society (the difference principle). (Rawls, 2001, pp. 42-3).

"Fair equality of opportunity" means that there should not be discriminating legal and social barriers that bar some sections of the society from accessing social institutions. The "difference principle" is the idea that inequalities would only be accepted if they somehow benefit the worst-off in society. For example, if cuts in both taxes and government spending ultimately benefit everyone by making us all richer, but at the cost of some loss of equality, then we would be willing to accept this *as long as it* really benefits everyone, including the poorest. On the other hand, an appeal for aggregate or average welfare would not be a rational choice because, under the veil of ignorance, such a deal does not ensure that the individual would end up in a group that is better off.

Is a Rawlsian framework appropriate for considering the distribution of education? Education has particular characteristics. It is both an end in itself – often considered as a basic right and a basis for self-respect – and a means to several ends, including economic gains and the ability to participate in a democratic society. It also has positional characteristics: competition in labour markets means that there are gains to being better qualified than the average worker. The distribution of educational opportunities required to ensure both fair equality of opportunity and the difference principle's effect on life chances more broadly, may be different – potentially more equal or redistributive – than that implied by simply removing discriminatory legal and social barriers to accessing education.

Rawls' formulation has come under criticism from many directions yet retains a central place in the philosophy of equity. The "veil of ignorance" is an attempt to explain the role of impartiality – of ignoring characteristics like race or wealth – in formulating principles for equity. The principles that emerge from this thought experiment, Rawls argues, go further than what a basic idea of equality of opportunity would allow. They allow for some inequalities between groups in outcomes, but not in basic liberties, nor in opportunities to access social institutions, and emphasise the need to remove both social and legal barriers to such opportunities.

Individual differences, capabilities and redistribution

The fairness of a distribution also depends on what, exactly, is being measured. Amartya Sen (2003) argues that when evaluating the quality of life, one has to move beyond "commodity fetishism", that is, a focus on the distribution of money or goods, and focus instead on evaluating the freedom that people have to lead the type of life they value (Sen, 1992). Sen distinguishes "functionings" or the things that people actually achieve from "capabilities", which are the set of functionings that people have open to them and are able to choose between. Sen's approach recognises that different people will have different goals or ends and argues that an evaluation framework for equity must take into account such differences (Sen, 2003).

Sen criticises Rawls' approach for focusing on goods – the means to achieving freedoms – rather than on the freedoms or capabilities themselves. The problem with this, in Sen's view, is that "people's ability to convert primary goods into achievements differs, so that an interpersonal comparison based on the holdings of primary goods cannot, in general, also reflect the ranking of their respective freedoms to pursue any given – or variable – ends" (2003, p. 48). The capability approach pushes us to consider, for example, the life opportunities that may be opened to someone through a given number of years in school, rather than the years in school themselves, or even the learning outcomes that result from the years in school. Similarly, a person who chooses not to attend higher education cannot be considered equally deprived as someone who has no such option, even though their observed educational attainment may be the same.[3]

3 The example Sen gives is that a person who deliberately fasts cannot be taken to be equally deprived as one who starves because of extreme poverty, even though on a superficial level, their functionings are the same.

The capability approach, however, poses particular difficulties for international measurement. For instance, we do not know *a priori* which kind of life different individuals value and which achievements are a matter of individual choice or social constraint, and so may end up using proxies in the form of goods (educational expenditure, school facilities) and achievements (learning outcomes). Also, with this approach more in-depth analysis is needed to place these proxies in the broader context of each individual's set of opportunities in life. In some contexts, for example, we may wish to consider the extent to which an education system can, or should, compensate for disadvantage due to disability or being born into a poor family in order to promote equality in capabilities.

Focus on the consequences of unequal education

There is an implicit assumption in both equality of opportunity and Rawlsian justice that societies have to make trade-offs between equality and other social goals, such as economic efficiency. Both approaches are trying to find a balance where some inequality exists but which can be justified morally and politically. However, this assumption needs to be questioned carefully in specific cases. Given a real country with a particular level of inequality in educational resources or learning outcomes, would more inequality really damage the economy or lower average welfare? And would less inequality actually improve economic growth or average welfare? But in either case, we would be less concerned with the question of what level of inequality can be justified and more concerned with reducing inequalities to avoid the problems they create.

The research on the consequences of educational inequalities, particularly in developing countries, remains limited but suggests that differences in access to education and wide disparities in learning outcomes are associated with a range of negative aggregate outcomes, including slower economic growth and higher risk of violent conflict. Concerning incomes, Wilkinson and Pickett (2010) report that income inequalities are associated with a wide range of undesirable outcomes, including in health, education and happiness. Research in OECD countries found that income inequality generally has a negative impact on economic growth (Cingano, 2014). The same study also reveals that increased income disparities depress skills development (e.g. years of schooling, skills proficiency) among individuals from low socio-economic backgrounds and suggests that education policy should focus on improving access for low-income groups.

Simply focusing on the public economic returns from education is sometimes sufficient to argue for a more equal distribution of educational resources. For example, primary education is often found to have higher economic returns than secondary or higher education; the boost in productivity as workers go from no education to primary education is higher than the boost associated with moving from primary education to secondary education. This means that targeting additional educational investments to those who do not currently complete primary education, a policy tending to equalise the distribution, would have the highest returns.[4]

The outcomes in one domain often influence opportunities in another. For example, income inequalities in many societies determine what access one has to educational opportunities. In order to equalise educational opportunities, policymakers could either try to break the link between income and education, for example by abolishing school fees, or focus on reducing income inequalities in the first place, for example through redistributive taxation.

Stewart (2002) argues that "horizontal" inequalities – inequalities between culturally-defined or constructed

4 Arrow (1971) is a classic treatment of the optimal allocation of educational expenditure from a utilitarian perspective.

groups, such as by ethnicity, religion, race, region and class – may have particularly negative consequences. The concept of horizontal inequalities has close affinities with equality of opportunity, when considering how outcomes may differ between groups in ways that are not morally or politically sustainable. But horizontal inequalities focus more closely on the particular nature of culturally-defined groups, as opposed to, for example, the differences between rich and poor. Inequalities between groups tend to persist or widen over time, because of the cumulative and mutually-reinforcing nature of disparities in economic, social, political and human capital (Stewart and Langer, 2008). As a result, horizontal inequalities can have long-term implications for stability and security in society (Stewart, 2002).

In addition, discrimination against a group may be economically inefficient because it holds back talented individuals within that group (Stewart, 2008). Cederman et al. (2011) and others have shown that persistence of group inequalities raises the risk of conflict and instability significantly. Group inequalities are powerful grievances that leaders can exploit to mobilise people for political protests. Civil conflicts are typically between groups and so arguably inter-group inequalities are most likely to be important in triggering conflict (Østby, 2008). Cross-country empirical analysis confirms a relationship between inter-ethnic inequalities in educational attainment and conflict (Østby and Urdal, 2010).

Although evidence of the effects of educational inequalities on other outcomes is still somewhat limited, the likely costs and trade-offs of more equal learning, as well as the likely benefits, need to be kept in mind when considering equity measures. Inequalities in learning outcomes may be socially, economically or morally problematic even if opportunities are in some sense equal. Decisions about what forms of inequality are acceptable or unacceptable need to be made in relation to particular country contexts and policy goals.

Spheres of justice: Applying different principles in different domains

The approaches described above tend to push for a single principle governing the distribution of a number of different educational goods. Are access to primary schools and good teachers, literacy and numeracy skills, university places and public educational expenditure all to be subjected to the same expectations when it comes to how these goods should ideally be distributed? This seems doubtful when we consider how these different goods serve different purposes and feed into each other. For example, an unequal distribution of expenditure may be needed to assure an equal distribution of literacy. Walzer's (1984) *Spheres of Justice* provides useful insights on this point, suggesting that we cannot expect a single distributional principle to apply to diverse goods. Walzer argues that the meaning of these different goods and their place in particular societies should determine the just distribution of each.

Walzer further suggests that the nature of the democratic state requires inclusive schools so that all citizens can grasp the body of knowledge needed to play one's part in a democratic society. All children have an equal need for the literacy and knowledge they can acquire in basic education, which derives from their equal membership as future citizens of a democratic state, and this membership "is best served if they are all taught the same things" (p. 202), regardless of the wealth or position of their parents. This vision of what basic education is for, if we agree with it, suggests that we might want to measure equity in basic education differently from education at higher levels. Our measures of equity in basic education would be based on the need for all children to reach a minimum standard and with limited variation in the learning outcomes they achieve. "Teaching children to read is," according to Walzer, "an egalitarian business, even if teaching literary criticism (say) is not" (p. 203).

Within schools, however, interest and capacity are important criteria for the distribution of knowledge, as well as need. Children's status as future citizens comes first, but once they have acquired what they need in this capacity, their status as future workers, managers or professionals can also come into play. Receiving a specialised education is perceived as being similar to "holding office" – a social position for which everyone has an equal right to be considered but which will ultimately be granted only to those who qualify. The presence of both "welfare" and "office" characteristics creates a deep strain in education, between the need for universal provision and the need to differentiate between students.

Walzer's argument draws on a broad-stroke sociological account of educational norms in democratic societies. The specifics of this account can be called into question by reference to empirical research on current, actual societies, but the central point would remain that distributional principles should be derived from such an account, rather than determined in advance in accordance with abstract universal principles:

> Welfare systems and markets, offices and families, schools and states are run on different principles: so they should be. The principles must somehow fit together within a single culture; they must be comprehensible across the different companies of men and women. But this doesn't rule out deep strains and odd juxtapositions. (Walzer, 1984, p. 318)

This way of thinking may be useful in considering how different distributions might be fair for different educational variables. For example, a meritocratic principle, similar to the distribution of "holding office" in Walzer's account, may apply to places in higher education, while a universalist principle would be more likely to be applied to the attainment of basic learning outcomes at primary level, which can be seen as fundamental for participation in society.

2.3 CONCEPTS FOR MEASURING EQUITY

How do we translate the philosophical debate on equality and equity into measurement of distributions in a data set? In this section, we present five key concepts that can be applied directly to a distribution. Their meaning in the broader equity debate depends on which indicators they are applied to. The concepts fit into two broad classes: some are "univariate", depending only on the distribution of some educational variable, while others are "bivariate" or "multivariate", depending on the joint distribution of education and one or more other characteristics such as wealth, gender, or parents' education *(see **Table 2.1**)*.

Table 2.1 Classification of equity concepts and related equity norms

Univariate based on the distribution of an educational variable	**Bivariate/multivariate** based on the joint distribution of an educational variable and one or more characteristics
Minimum standards binary educational variable (e.g. completed primary education) is positive for everyone	**Impartiality** education does not depend on background characteristics
Equality of condition educational variable is the same for everyone	**Meritocracy** education is positively related to ability but not related to other characteristics
	Redistribution education is positively related to disadvantage

Source: Authors' analysis.

2.3.1 Meritocracy

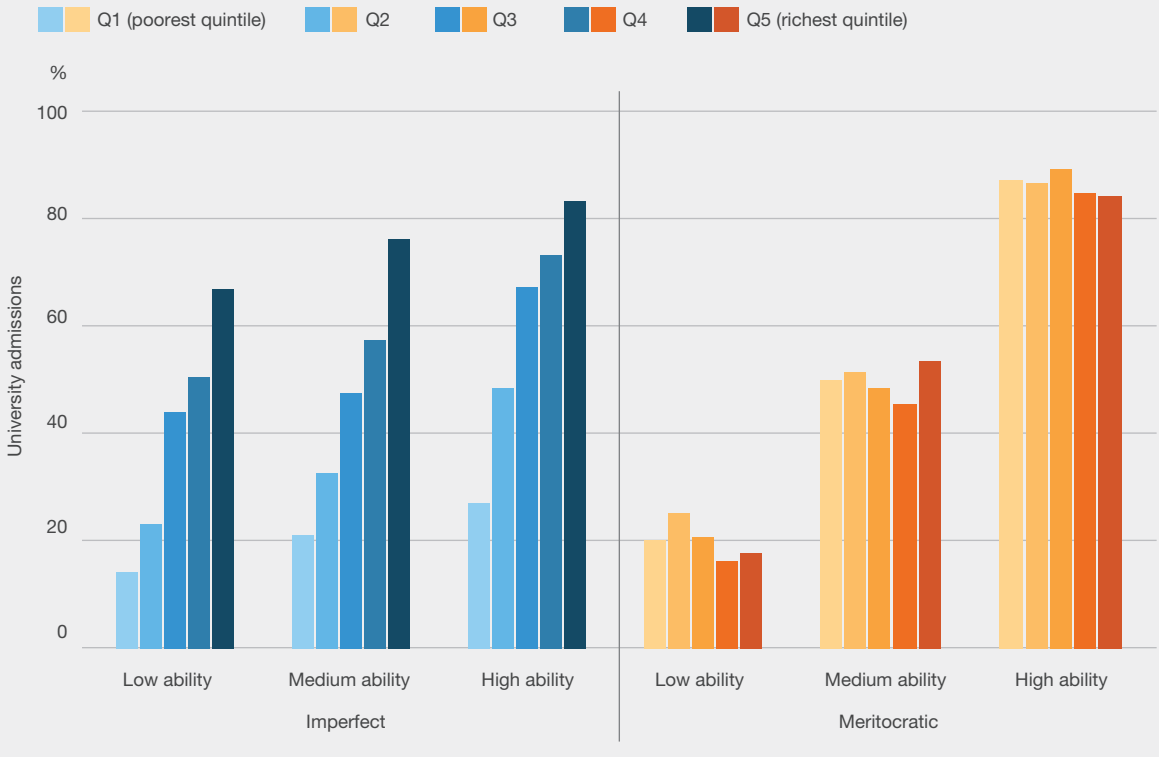
Meritocracy means that educational opportunities are distributed on the basis of merit. Many education systems apply *de facto* meritocratic principles to the distribution of educational opportunities. Children judged the most able, usually through performance in

high-stakes examinations at the end of each level of education, are given opportunities to continue through the system or given opportunities in a different type of education (e.g. academic vs. vocational) compared to their peers. Meritocracy means distributing education *unequally* with respect to a particular *relevant* difference reflecting individual merit. In practice, merit could mean intelligence, effort, accomplishment or some combination of these and may be measured through tests, references, etc. Meritocracy also implies that education will be distributed *equally* with respect to other, irrelevant differences.

Measurement of meritocracy requires adequate measures of the relevant form of merit, which may sometimes be contested. For example, exam scores may be used to measure a student's suitability for entrance to secondary or higher education, but those exam scores will not always perform well as a guide to the student's real ability.

The extent to which a system is meritocratic can be seen by examining whether the outcome of interest (e.g. university admissions in **Figure 2.1**) correlates with the measure of merit (e.g. academic ability as measured through test scores in upper secondary education), while being uncorrelated with supposedly irrelevant differences (e.g. wealth). In practice, however, the way opportunities are distributed and justified through meritocracy is often a source of controversy. For example, if ability in secondary education (measured through test scores) is correlated with wealth, then many systems that claim to be meritocratic will result in wealthier students

Figure 2.1 Imperfect and perfectly meritocratic distribution of university admissions



Note: Hypothetical data.

enjoying better opportunities than poorer ones. Such situations tend to generate political discussion about whether the type of merit being measured is really a fair basis for distributing opportunities.

Walzer's *Spheres of Justice* provides insights into how meritocratic principles may co-exist alongside more egalitarian principles within a single education system, in tension with each other but driven by different needs. For example, universal access to basic education may be driven by an egalitarian concern for an inclusive society or by a rights framework, while opportunities in post-basic education may be driven more by a concern for developing individuals who can become experts in particular fields, in line with their future working lives.

2.3.2 Minimum standards

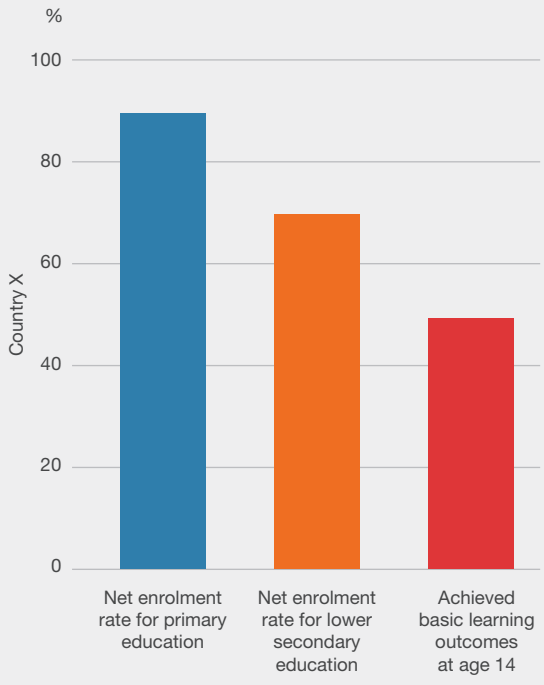
Many societies distribute educational opportunities on a meritocratic basis at the higher levels of education, while maintaining minimum standards in lower levels of the education system. For example, completion of primary school has long been seen as a right in many countries and the Sustainable Development Goals (SDGs) that were adopted by the United Nations in 2015 call for universal primary and secondary education.

The minimum standards principle involves seeing education in terms of a binary criterion – a child is enrolled in primary school or not, or can demonstrate basic literacy or not – and insisting that this criterion should be fulfilled for all individuals. Often, the minimum standard reflects a right or agreed norm. Simply measuring the proportion of individuals who meet the minimum standard could be taken as an equity measure – equity is achieved when 100% of individuals meet the standard. The Millennium Development Goals (MDGs) adopted by the United Nations in 2000 helped to establish the widespread use of indicators such as the net enrolment rate to measure how close countries are to meeting a standard of universal primary education. These measures can be visualised with simple and familiar charts showing how close countries are to 100% *(see **Figure 2.2**).* The new focus on equity in the SDGs and elsewhere means going beyond this type of analysis and looking at impartiality *(see Section 2.3.3)* in the proportions of individuals meeting minimum standards or the probability of an individual meeting a standard.

2.3.3 Impartiality

Equality of opportunity (described in Section 2.2.2) has become a dominant concept in normative frameworks for equity in education. It argues that educational goods should be distributed equally with respect to differences which should be irrelevant, such as gender, race, wealth or location. As noted, however, it is not clear that the philosophical concept

Figure 2.2 Proportion of children reaching three types of minimum standards



Note: Hypothetical data.

of equality of opportunity is the right term for this concept in education. Equality of opportunity implies holding people responsible for the things within their control and not for circumstances beyond their control; but the moral basis for holding children responsible for their own innate talent or motivation appears weak. For this reason, we refer to this type of equity concept in education as *impartiality.* We use this term to separate the moral from the political philosophy issue of ensuring that individuals have equal opportunities and from the statistical exercise of examining the extent to which a distribution depends on circumstances. Impartiality is centrally important in education, whether or not it is taken to represent equality of opportunity.

Impartiality is similar to the concept of horizontal equity in Stewart (2002) and the concept of equality of opportunity discussed in Berne and Stiefel (1984). One way of seeing it is to argue that the statistical measure of impartiality indicates a lower bound on true equality of opportunity. Not all circumstances are captured, but the easily measured ones are.[5] A focus on impartiality may also be justified by thinking of schools as a "sphere of justice", with aims that include ensuring all children reach a minimum standard and encouraging all children to learn to the best of their ability. It is inevitable, because of the way children learn, that differences will arise as a result of their different sets of abilities, interests and motivation. Policies aiming to compensate for differences in student motivation, for example, could end up demotivating the most enthusiastic learners. But it is not inevitable that differences will emerge as a result of, say, wealth or gender, as such differences are incongruous with the social role of the school. Against this, it might be argued that schools (and perhaps, society at large) should be held responsible for motivating children and should provide extra assistance to those who struggle the most. In this case, impartiality measures again only provide a lower bound on the extent of morally significant inequality.

Impartiality provides a way of checking that minimum standards are being equally met across different population groups, and of ensuring that an outwardly meritocratic system is not simply used to justify and entrench an unfair distribution of opportunities. Impartiality is also important because, on the one hand, rights frameworks insist that the education system should be free of discrimination and that different population groups should have an equal chance of accessing each type of opportunity; but on the other hand, they sometimes concede that (perhaps for reasons of insufficient supply, or differences in inherent ability) not everyone will have access to every level of the education system. Moreover, impartiality measures are important because they can point us directly towards the most disadvantaged groups who can then be targeted by policy.

Impartiality measures essentially quantify the relationship between an education indicator of interest and one or more measures of circumstance, and define perfect impartiality as the absence of any relationship. Analysis of impartiality is value-laden, because we have to select which characteristics to count as circumstances, and which to see as legitimate sources of variation.

Impartiality measures can be grouped into five main types *(see **Table 2.2**).* In many cases, the easiest and most accessible analysis of impartiality involves simply presenting statistics disaggregated by different groups in a table or graph *(see the example in **Figure 2.3**).* Tabulating the **gaps or differences** between particular groups, such as the difference between the richest and poorest, enables comparisons to be made across countries or over time.

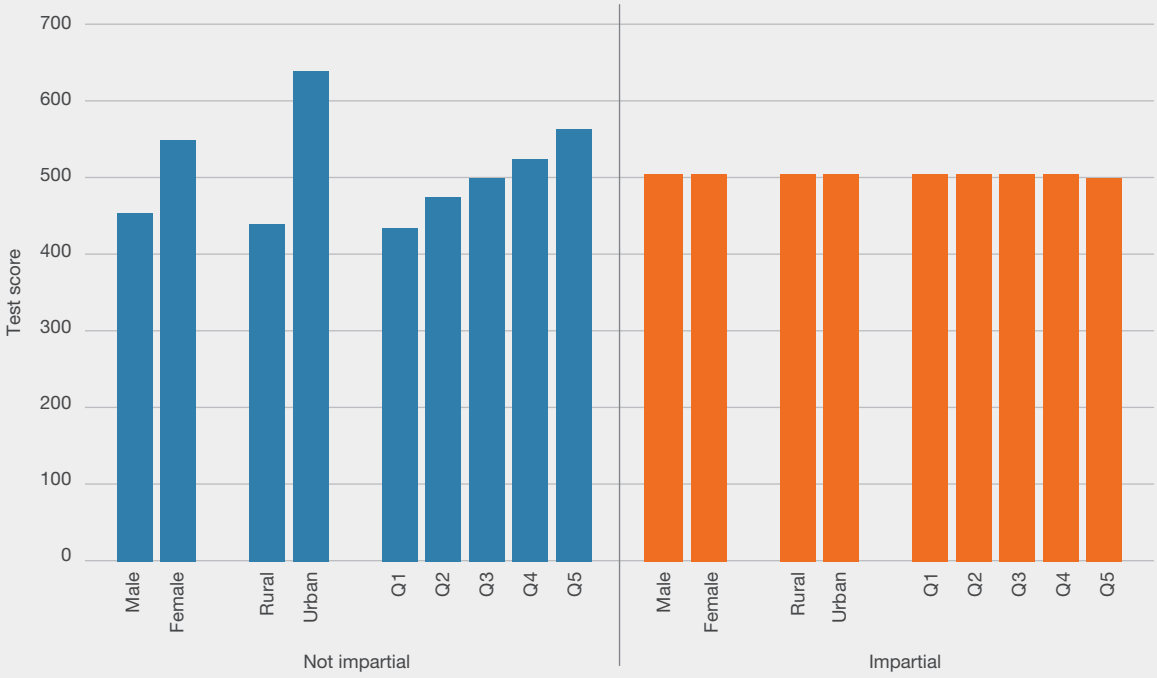
5 As Ferreira and Gignoux say of their "inequality of opportunity" measure: "It is a parametric approximation to the lower bound on the share of overall inequality in educational achievement that is causally explained by pre-determined circumstances" (2011, p. 17).

Table 2.2 Selected impartiality measures

Family of metrics	Impartiality measures and analyses	Description and remarks
Gap, difference	Cross-tabulation or disaggregated bar charts	Show the value of an indicator for each group. Simple but very effective means of visualising impartiality. More complicated charts (like those used in UNESCO's WIDE database) can visualise multiple overlapping sources of deprivation.
Ratio	Parity indices	The gender parity index is the ratio of female to male values of a given indicator. Similar parity indices can be calculated for wealth quintiles, rural and urban location, etc.
Co-variation	Correlation coefficient (r)	Strength of linear relationship between two variables, ranging from -1 (perfectly negatively related) to +1 (perfectly positively related).
	Slope index of inequality (SII)	Coefficient in an ordinary least squares (OLS) or weighted least squares (WLS) regression, e.g. of years of education on wealth (Wagstaff et al., 1991). Individual or grouped (e.g. district or wealth quintile) data may be used.
	Relative index of inequality (RII)	Slope index of inequality divided by the mean level of the dependent (education) variable (Wagstaff et al., 1991).
	Elasticity	The percentage effect of a 1% change in the characteristic (e.g. poverty) on the indicator (e.g. government education expenditure). Elasticity can be calculated from regression coefficients (see Berne and Stiefel, 1984, Table 2.4).
	Proportion of variance explained by circumstances (R^2)	Uses OLS regression to examine the proportion of the variance in an indicator explained by circumstances like wealth and location. Variants use different regression specifications, or a measure of total inequality other than the variance.
	Ordinal segregation	Measure of the ratio of between-category variation to total variation (see Reardon, 2009).
	Dissimilarity index (D)	Weighted average of the gap in probability of access to education between different circumstance groups and the overall average access rate (see Paes de Barros et al., 2008).
Concentration	Concentration curve and concentration index	Analogous to the Lorenz curve and Gini coefficient (see Table 2.3). The concentration curve plots the cumulative proportions of the population by wealth (starting with the most disadvantaged) against the cumulative proportions of education. The concentration index is twice the area between the curve and the diagonal line that represents perfect equality. Used extensively in the health inequality literature (see Erreygers and van Ourti, 2012).
Group-level cumulative information	Group standard deviation, group Gini	Measures of equality of condition, such as the standard deviation, can also be applied at the group level (e.g. between districts or ethnic groups), and interpreted as a measure of impartiality between the groups.

Source: Authors' analysis.

Figure 2.3 Impartiality of education based on test scores by sex, location and wealth quintile



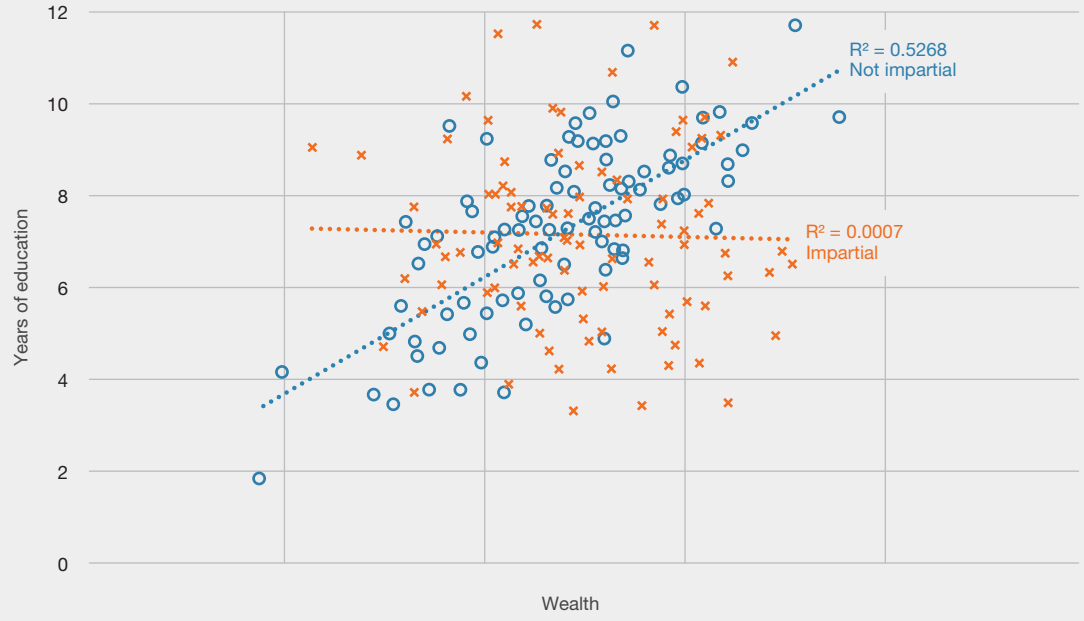
Note: Hypothetical data.

In some instances, **ratios** are a more appropriate, but still simple, way of showing differences between groups. Parity indices are a common measure used especially to measure gender parity in education but only work when the characteristic in question is a single binary variable. Parity with respect to wealth can be calculated by comparing the richest 50% to the poorest 50%, or by ignoring the middle of the wealth distribution and comparing, for example, the richest 20% to the poorest 20%. Parity indices are among the official SDG indicators, to be applied to all education indicators for female/male, rural/urban, poorest/richest wealth quintiles and other dimensions as they become available (United Nations, 2017).

For continuous variables, analysts commonly turn to measures of **co-variation,** such as correlation or regression coefficients. To understand impartiality with respect to multiple variables, we need a measure such as the coefficient of determination (R^2), representing the proportion of the total dispersion explained by those variables. For example, when wealth is plotted against years of education *(see **Figure 2.4**),* a coefficient of determination close to zero suggests that there is no relationship and that educational attainment is impartial with respect to wealth. A high R^2 would suggest that educational attainment is strongly correlated with wealth and so is not impartial.

There is an important distinction between measures of the strength of relationship, such as the correlation coefficient, and measures of the magnitude of the relationship, such as the regression slope. A high-magnitude but low-strength relationship means that we cannot consistently predict, say, educational

Figure 2.4 Impartiality of education with respect to wealth



Note: Hypothetical data.

outcomes given an individual's wealth, but that our best estimate of the relationship suggests that large differences are likely. A low-magnitude, high-strength relationship means that the difference between rich and poor is not that large but is very consistent. Both of these concepts may be important and it may be useful to present both types of measure *(see discussion in Berne and Stiefel, 1984)*.

Paes de Barros et al. (2008) create a Human Opportunity Index, which combines the overall rate of access to education with the *dissimilarity index* for access to education. The dissimilarity index is defined as a weighted average of the gaps between the probability of accessing education by different population groups and the overall average rate. The Human Opportunity Index can thus be seen as a combined measure of minimum standards (the overall rate of access) and impartiality (the dissimilarity index).

Concentration curves and **concentration indices** provide an analogy to the Gini coefficient *(see Section 2.3.4).* These are often used in assessing health equity but rarely applied to education, despite being a potentially useful tool for examining the joint distribution of an education variable with respect to a second variable, such as income.

Finally, measures such as the Gini coefficient or standard deviation – which, when applied at the individual level, measure equality of condition – become impartiality measures when applied at the group level. For example, the standard deviation of average education expenditure among the districts in a country provides a measure of impartiality between the districts.

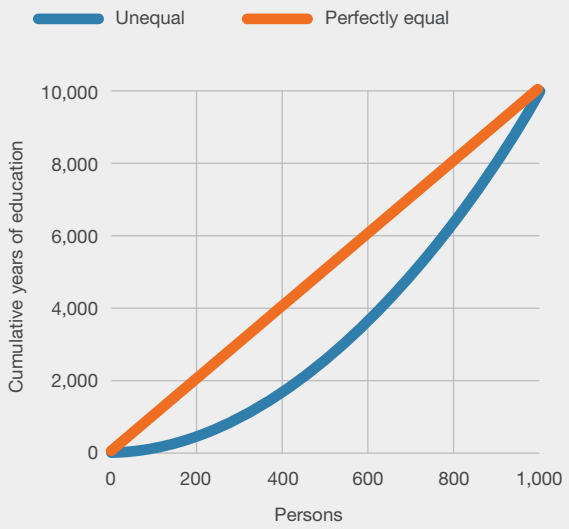
Should impartiality measures control for prior ability? Often it will be of interest to know whether students' progression – say, the degree to which their cognitive abilities have improved between the beginning and

end of primary school – is impartial with respect to wealth or gender. Their abilities upon entering primary school are the result of their innate ability, the early childhood learning environment and pre-primary education. Depending on what policy question is being asked, we might be interested in how much students progress during primary school, controlling for their starting position, or we might be more interested in where they end up at the end of primary school, regardless of where they started. Impartiality conditioned on ability becomes similar to measures of meritocracy *(see Section 2.3.1),* depending on what measure of merit is considered a relevant one for the distribution of opportunities or outcomes.

2.3.4 Equality of condition

Impartiality is concerned with the way that an educational variable relates to circumstances such as wealth, gender or location. But another approach to measuring inequality is simply to look at the distribution of the educational variable across persons, regardless of their different circumstances. We refer to this as "equality of condition". For example, consider the number of years of education that each person in a population has. Starting with the person with the least education and ending with the person with the most education, we can plot the cumulative years of education against population *(see **Figure 2.5**).* This is the Lorenz curve. It is the basis for much analysis of economic inequality and can be applied to educational inequality too, provided we have a continuous educational variable to measure the amount of education each person has. A straight line represents perfect equality: everyone has the same number of years of education. The more curved the line, the more unequal the population is with regard to education.

Figure 2.5 Lorenz curves showing a perfectly equal, and somewhat unequal, distribution of years of education



Note: Hypothetical data.

Why would we be interested in equality of condition rather than impartiality? Certain educational inputs, goods or outcome thresholds should be distributed universally and equally, so that these are at the same level for every individual, regardless of whether we look within population groups or across groups. For example, it might be argued that public expenditure on education per student should be the same for every child. In education finance, Berne and Stiefel (1984) use the term "horizontal equity" to refer to the principle of "equal treatment of equals", noting that there is often an expectation that all students will benefit from equal levels of government expenditure. From a rights or citizenship perspective, for example, children may be seen as all equal and therefore all deserving of equal treatment.

Equality of condition measures are sometimes referred to as the "classic" or univariate inequality indices. Any measure of statistical dispersion, including the long list of indicators developed in the income inequality literature (Cowell, 2011; Atkinson, 1970; Theil, 1967), can be used *(see **Table 2.3**).* However, as we will discuss in Section 2.5.1, some are clearly better than others and some are more suited to specific tasks. The range measure, for instance, is a simple measure which is easily interpreted and makes

Table 2.3 Selected equality of condition measures

Family of metrics	Equality of condition measures and analyses	Description and remarks
Difference, gap	Range	Difference between highest and lowest indicator, across individuals, schools, districts, etc.
	Restricted range	Difference in the indicator at specific percentiles in the distribution, e.g. interquartile range is the difference between the 75th and 25th percentiles.
Ratio	Palma ratio	Ratio of the share of education of the top 10% of the distribution to that of the bottom 40%. Variants use other percentiles. Considered policy-relevant for income but not yet applied in education.
Dispersion	Variance or standard deviation	Average squared deviation (difference from the mean) in the indicator.
	Coefficient of variation	Standard deviation divided by the mean.
	Mean absolute deviation	Average absolute deviation (difference from the mean) in the indicator.
Concentration, cumulative information	Atkinson index	Based on the choice of an explicit social-welfare function (Atkinson, 1970).
	McLoone index	Sum of the indicator for individuals below the median, divided by the sum of the indicator if all indicators were at the median.
	Lorenz curve and Gini coefficient	Relationship between the actual distribution and a uniform distribution.
	Theil index	Based on the notion of entropy in information (Theil, 1967).

Note: Parts of this table are based on Table 1 in Toutkoushian and Michael (2007) and on the list presented in Berne and Stiefel (1984, p. 19).

minimal demands on the data but lacks the important property of being sensitive to changes in the middle of the distribution. The Gini coefficient, commonly used in the income inequality literature, is more sensitive to changes in the middle of the distribution than at the top or bottom.[6] Many indicators taken from the income or health inequality literature are *relative* measures, while in education we may often be more interested in absolute inequality measures *(see Section 2.5.1).* Other indicators deliberately focus on specific parts of the distribution; for example, the McLoone index focuses on individuals whose education is below the median and so is relevant in a policy context where the objective is to bring everyone up to a given level. Equality of condition measures require a continuous variable. The minimum standards measures discussed in Section 2.3.2 can be seen as a way of looking at equality of condition when we have a binary variable; we simply track what proportion of individuals meet the minimum standard. Nominal and ordinal variables *(see Section 2.4)* offer no obvious way of measuring equality of condition, unless they are converted into interval or ratio scales, which could be done by converting the grade someone has completed at school into a number

6 The Gini coefficient is a measure of inequality among individuals in a group, with values that can range from 0 (indicating that everyone is equal) to 1 (indicating perfect inequality, with one person earning all income, for example).

of years of education or by converting examination grades into points.

The meaning of equality of condition also depends on the unit of analysis. Analysis may be at the level of the child, school, district, province, state or country. Measuring equality of condition with provincial-level data is similar to measuring impartiality between provinces, although in practice the measures may differ, depending, for example, on whether province-level analysis is weighted by the number of children or students in each province.[7]

Some analysts object to equality of condition measures on the basis that equality is not the same as equity (a fair distribution) or that total equality in something like learning outcomes will never be achieved simply because individuals are different or because we lack any benchmark for what level of equality (for example, what value of the Gini coefficient) we should be aiming for. These criticisms all contain some truth, although it is possible in principle to estimate an optimal level of equality with regard to some specific policy goal. For example, one could use cross-country regression to estimate what value of the Gini is associated with fastest economic growth (or fastest progress towards achieving some other social goal).

Often, impartiality measures will capture most of the inequality of condition: much inequality is between groups such as rich and poor, and the part that remains within groups may be both small and difficult to eradicate because it relates to unobserved differences such as innate ability or work preference. It remains important to consider the total dispersion of key indicators and the trends in this dispersion over time. Univariate measures also have the advantage of requiring relatively little data (by definition, only one variable is required), and of providing equity indicators that are comparable over time and between countries.

2.3.5 Redistribution

In order to move towards impartiality or equality of condition in educational outcomes, governments may choose to distribute educational inputs unequally, in ways that compensate for existing disadvantage. *Redistribution* indicators are of particular interest in the field of education finance. They can measure the extent to which the distribution of some educational variable, e.g. public education expenditure, compensates for some degree of existing disadvantage, such as regional poverty rates.

Governments sometimes allocate more public spending to historically-disadvantaged regions, in an attempt to equalise learning outcomes. Berne and Stiefel (1984) refer to this as "vertical equity," which they define as "appropriately unequal treatment of unequals", as opposed to "horizontal equity," which, as mentioned above, refers to the equal treatment of equals. The idea also echoes Sen's (1999; 2002) concern that an equal distribution of goods does not necessarily translate into an equal distribution of functionings or freedoms. Children with learning disabilities or whose school uses a language other than their mother tongue, for example, may merit appropriately unequal treatment requiring more expenditure or other resources.

In school financing, some districts may face higher costs than others because their schools are more remote, higher salaries need to be offered to attract equally-capable teachers, or students need more specialist teachers to reach an equal level of learning. A national financing mechanism that gives equal per-student funding to each district would ignore these extra needs and costs. As an illustration of a redistributive or vertical equity analysis, Berne and Stiefel (1994) use regression to look at the relationship between resources (such as expenditure) and poverty, by sub-district or school, in New York City in the 1990s. They find *higher* per-pupil expenditures in sub-

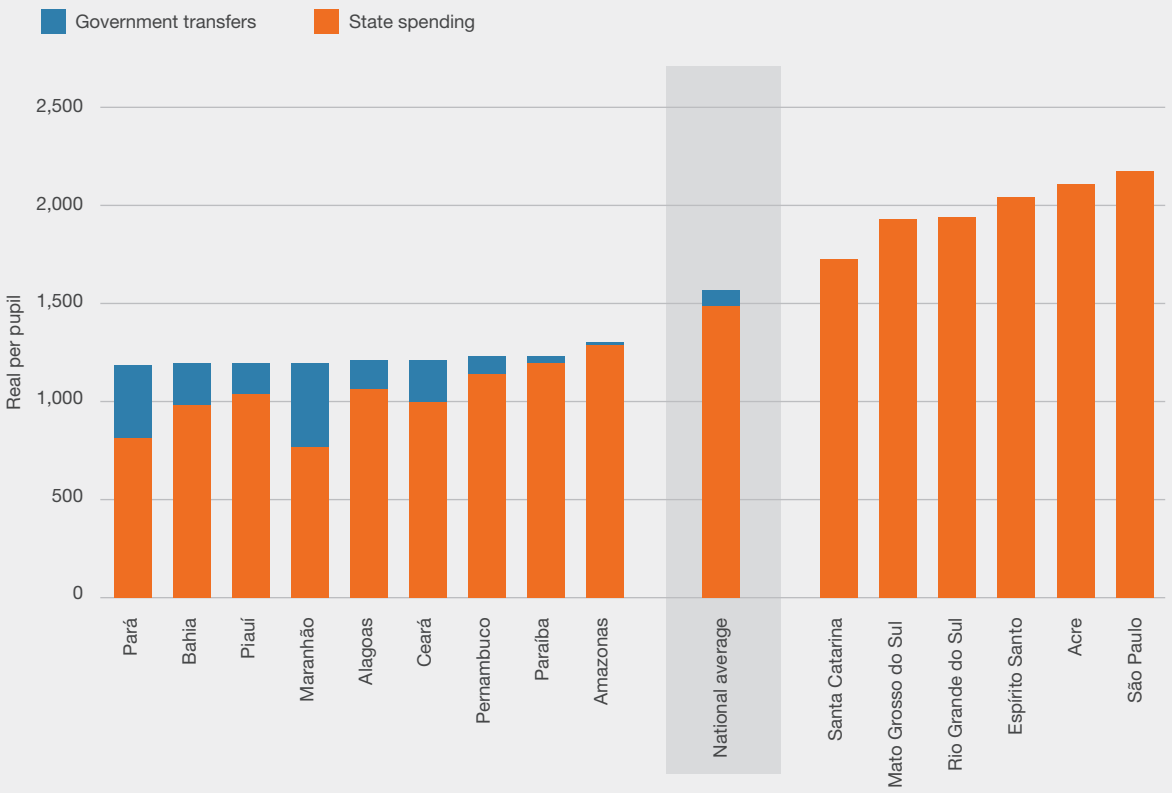
7 See Berne and Stiefel (1984) for further discussion of pupil- and district-level analysis.

districts with *lower* poverty, suggesting a regressive distribution in elementary schools.

Countries with federal government structures may also practice redistributive financing. In Brazil, for example, complementary federal financing tops up state education spending for states whose tax revenues leave them below a stipulated threshold (UNESCO, 2010; **Figure 2.6**). This eliminates inequality below the threshold level, but leaves large inequalities in spending per pupil across the country as a whole.

There is much controversy over which characteristics merit differential treatment in a redistributive approach. In the United States, for example, Berne and Stiefel (1984) note that learning disabilities and speaking English as a second language are often seen as meriting higher levels of education expenditure, while characteristics such as race and gender are not. Many education systems show much higher expenditures at higher levels of the education system (see e.g. UIS, 2016), even though many educationalists would point out that early childhood is the most important stage and thus deserves at least a larger share of the expenditure than it currently gets. They are supported by the additional argument for redistribution which holds that, due to drop-out, children from the most disadvantaged backgrounds are often enrolled in much larger numbers in lower rather than higher grades.

Figure 2.6 Federal government redistribution of education financing in Brazil



Source: Taken from UNESCO (2010); Henriques (2009), based on data from Fundeb.

Table 2.4 Selected redistributive equity measures

Measure or analysis	Description and remarks
Weighted dispersion measures	Observations are weighted based on the inverse of some characteristic (e.g. poverty rates in each district), and then equality of condition measures (e.g. variance) are applied.
Ratio analysis	Ratio of indicator (e.g. government education expenditure) across groups (e.g. low- and high-income districts).
Regression slope	The effect of a unit change in the characteristic under consideration (e.g. poverty) on the indicator (e.g. government education expenditure). Simple correlations can also be used.
Elasticity	The percentage effect of a 1% change in the characteristic (e.g. poverty) on the indicator (e.g. government education expenditure).

Note: Parts of this table are based on Table 1 in Toutkoushian and Michael (2007).

Table 2.4 lists some common measures of redistributive equity. There are two distinct approaches to measurement of redistributive equity (Berne and Stiefel, 1984). The first starts with a specific view on how much inequality between unequals would be acceptable. For example, a policymaker might decide after a review of evidence, that children from poor backgrounds need to be allocated twice as much public education expenditure as children from rich backgrounds if they are to have a fair chance in the education system. It is then possible to identify with a single statistic the extent to which the actual distribution resembles the desired distribution. The population of children can be weighted according to their need and the Gini coefficient calculated on education expenditure per child for the weighted data. A federal government that aims to equalise expenditure across states serves as another example: simply looking at the sum of federal and state expenditure and applying a univariate inequality measure, such as the Gini or variance, would measure the proximity of the actual distribution to the targeted one. In both of these examples, a higher value of the measure means that the distribution is closer to our idea of a fair distribution.

In the second approach, we do not have any specific notion of how much redistribution would be appropriate but want to measure how much redistribution has taken place. This can be done using measures such as the regression slope or elasticity. However, we cannot always assume in such a case that a higher value for the measure is fairer, because there may be a point where more redistribution is taking place than is desired.

Benefit incidence analysis (Lassibille and Tan, 2007) is closely related to redistributive analysis. By examining the enrolment rates for different population groups (e.g. rich and poor) in different levels of education (e.g. primary and secondary) and the amount of government expenditure for each level of education, the amount of government expenditure per student in each population group can be estimated. This allows us to understand whether the poorest, for example, are receiving a fair share of government spending on education or if spending benefits the richest most (which can happen because in higher levels of education a larger proportion of students tends to be from wealthier households).

2.4 EQUALITY OF WHAT?

The preceding section described five different ways of measuring or analysing the distribution of an education indicator. In this section we consider the range of indicators that are used in educational analysis and suggest that indicators can be classified according to two dimensions that particularly matter for equity measurement. The first dimension is the stage in the education production cycle that the indicator relates to: is it an input, such as public expenditure; an outcome measure, such as children's

scores in literacy assessments; or something in between, such as school quality, teacher qualifications or enrolment? The second dimension concerns the type of variable: its level of measurement and whether it is bounded or not.

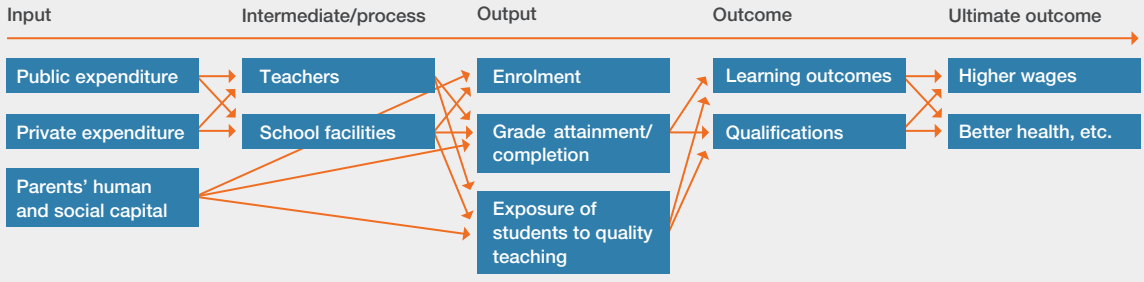
Figure 2.7 presents a simple model of how an education system works, which we can use to help classify education indicators. "Inputs" to the system include public (government) expenditures and private (household) expenditures. To this could be added expenditures by donors, non-governmental organizations (NGOs) and charities, and public expenditure could be broken down further by level of government (for example, federal, state and local). The non-financial support that parents and, perhaps, the wider community provide to children also feeds into the education production process.

These investments, if well-directed, are used to create school facilities and hire teachers, which can be classed as "intermediate" or "process" inputs. Indicators, such as the pupil-teacher ratio and an index of school facilities, fit in this category. They are the immediate outputs of one set of processes: investment in educational facilities and human resources; and the inputs into another set of processes: the learning that happens within schools.

The time students spend in school, measured in terms of enrolment, attendance, grade attainment and completion, are shown in Figure 2.7 as outputs of the education system. They are the results of a prior process: the construction and maintenance of well-functioning, accessible schools necessary for learning. Low attendance and high drop-out rates may be consequences of poor school quality (Sabates et al., 2010; Hunt, 2008). However, these outputs can equally be seen as inputs into the learning process. Teachers' time and the quality of school facilities are combined with students' time in school to produce learning. In the absence of data on learning outcomes, measures such as grade attainment, enrolment or completion rates have often been used as the main indicator of educational outcomes. But these proxies are known to be flawed because there may be large differences in quality across countries and over time, so that six years of schooling is associated with very different levels of learning in different contexts (e.g. Ferreira and Gignoux, 2011). Using grade attainment as an outcome indicator may be valid within countries or regions where there is relatively little variation in school quality, but these indicators should generally be seen either as educational outputs or as proxies for educational outcomes rather than as direct measures.

In simple terms, the main outcome of an education system is the amount of learning that has taken

Figure 2.7 A model of how an education system works



Source: Authors' analysis.

Box 2.1 Defining learning

There are a great variety of theories regarding the concept of learning (Illeris, 2009), emerging from the fields of educational psychology, neuropsychology, learning theory and pedagogy. Some theories are complementary while others are strongly contested. In order to avoid complex and contested multi-disciplinary theoretical landscapes, we employ a broad definition of learning, incorporating those elements which are generally shared across theories.

Illeris (2007, p. 3) states that "learning can broadly be defined as any process that in living organisms leads to permanent capacity change and which is not solely due to biological maturation or ageing". Therefore, learning is not just about the development of cognitive skills but also refers to the development of emotional and social skills and the interaction with the learner's environment. Learning "involves ongoing, active processes of inquiry, engagement and participation in the world around us" (Friesen et al., 2015, p. 5, citing Bransford et al., 2000). Dumont, Istance, and Benavides (2010, p. 17) state that the learning environment promotes "horizontal connectedness across areas of knowledge and subject as well as to the community and the wider world". These dimensions of learning are inextricably intertwined through two basic processes. The first is external, whereby the learner engages with the social, cultural or material environment, and the second is internal, whereby the learner's internal thought processes elaborate information and acquire capacity (Illeris, 2007). Learning draws on cognitive, emotional and biological resources in that learning relies on "meta-cognitive skills" (for learners to monitor and evaluate their own learning), it relies on learners being able to regulate their own emotions and motivations to learn (Dumont et al., 2010, p. 17), and it relies on social processes whereby knowledge is constructed through interaction and cooperation (ibid., p. 15).

These external and internal processes are adaptive and integrated. For example, the brain is both adaptive to the environment – the structure of the brain changes in response to physical events – while also influencing the effects of subsequent experiences (Hinton, 2005). Prior learning is one of the most important resources from which learners draw during the learning process. Therefore, prior knowledge, interest, motivations, self-efficacy, beliefs and emotions and linguistic, cultural and social backgrounds contribute significantly to the learning process (Dumont et al. 2010).

place. This handbook proposes a broad definition of learning incorporating ideas that are common across different learning theories *(see **Box 2.1**).* These all share the notion that learning can be measured using standardised international assessment instruments or using national examination systems. Moreover, they can be related to the expectations embodied in the curriculum of proficiency at each grade or to broader and less country-specific benchmarks.

What is currently measured tends to focus on a relatively narrow sub-set of all the forms of learning that could be measured, consisting mainly of literacy and language skills, mathematics and science. **Annex A** lists the major international learning assessment exercises that are currently in use, their geographic coverage, the learning domains they cover, and their comparability across countries.

Learning assessments demonstrate different strengths and weaknesses when measuring equity. For example, they deal in different ways with minority languages, have sample sizes that allow for different types of disaggregation, and some have "floor" or "ceiling" effects that make them less sensitive when it comes to distinguishing between the weakest or strongest learners. The annex also lists some of the main equity concerns arising from each of the current international assessment exercises.

Measuring learning outcomes usually means choosing a grade or age and administering a standardised assessment to a sample or to all individuals who fall into the target group. This may not be the best way. It is also possible to measure the progress that a student makes between two points in time, such as between the beginning and end of lower secondary

school. This would provide a better measure of the learning gain that has been achieved through the education system, although it would still also reflect the influence of the home background.

This handbook is mainly concerned with learning. However, it is important to remember that learning is also a means to an end. The ultimate outcomes of education include higher productivity, higher incomes, better employment prospects and various other desirable outcomes at the individual and social levels. From the point of view of individuals and their families, qualifications – and the learning achievements they represent – are also an important outcome of education. Educational qualifications are often used to determine who has access to ultimate outcomes, such as better-paid jobs. Individuals may also derive non-learning benefits from school, such as enjoyment, self-discipline, protection from hazardous child labour or the forming of social networks. Unfortunately, the full range of valued outcomes of education are rarely measured.

The level of measurement (Stevens, 1946; Roberts, 1979) is also important for the type of equity measure that we can apply to an indicator *(see **Table 2.5**).* Most inequality measurement is concerned with *interval* or *ratio* level variables, which can be measured on a continuous scale. These include variables such as assessment results, grade attainment and indices of school facilities or teacher effectiveness.[8] Some inequality indicators can only be calculated for ratio level variables, which have a meaningful zero point and where the ratio between the values for two individuals is meaningful but not for interval level variables, which do not have these properties.

It is generally harder to measure inequality in terms of a categorical (nominal or ordinal) variable.

Table 2.5 Levels of measurement

Level of measurement	Description	Example
Nominal	Unordered categories: individuals can be classified but not ordered.	Schools may sort students into arts and science streams. As long as neither stream is considered "better" than the other, the two streams constitute nominal categories.
Ordinal	Ordered categories: individuals can be ordered, but the differences between individuals are meaningless.	Schools sort individuals into top, middle and bottom sets by ability. The top set is better than the middle and the middle set is better than the bottom, but we cannot assign meaningful numbers to the sets.
Interval (cardinal)	Continuous but with an arbitrary zero point: differences between individuals mean something, but ratios do not.	A normalised index of school facilities.
Ratio	Continuous: ratios between individuals mean something; zero point corresponds to a complete absence of something.	Years of schooling, considered as if this is a continuous variable.

Source: Roberts (1979); Erreygers and van Ourti (2011).

8 Strictly speaking, grade attainment is a discrete variable – it can only take certain values from a finite set. Still, it can be treated as a continuous measure (length of time spent in education, measured to the nearest year) for many purposes.

Nonetheless, a categorical variable at the individual level, such as whether the individual is enrolled in school or not, can be turned into a ratio level variable at the group level, by taking the percentage of those who fall into a particular category within the group (e.g. the net enrolment rate within a region or wealth quintile). In most cases, the categorical variables of interest are ordered binary variables, such as enrolment or attendance, where one category (enrolled/attending) is clearly preferable to the other (not enrolled/not attending).[9]

Nominal variables – categories which do not have any particular order – may also be of interest for measuring educational inequalities, although this handbook does not focus on them. It is possible, for example, to examine impartiality in unordered categories, such as whether students from poorer backgrounds are more likely to enter vocational as opposed to academic streams. Even if the two streams are considered equally good, the excessive segregation of one group into one stream is likely to be problematic, for instance because of peer effects or the types of social networks that are built during schooling. The unequal distribution of groups among unordered categories is referred to as segregation (Massey and Denton, 1988; James and Taeuber, 1985) and can be analysed using a set of metrics similar to those used for measuring inequalities. These can be seen as measures of impartiality.

Variables can be either bounded or unbounded. Many variables are percentages which can only take values between 0 and 100%, meaning that they are bounded at both ends. This has implications for the types of inequality measure that can be applied.

Chapter 3 considers in more detail how these aspects of different educational variables affect the types of equity measure that we can use.

2.5 WHAT SHOULD AN EQUITY MEASURE LOOK LIKE?

A large number of inequality indicators exist, and these are well represented in the literature on income and health inequality. A comprehensive discussion of them is beyond the scope of this handbook, but it is useful to consider what characteristics they should have in different circumstances and with different indicators.

There is at least one property that is desirable for all kinds of inequality indicators. This is *population independence.* In general, an increase in the population (of a country, state or district), while retaining the same shape of the distribution of the indicators of interest, should not change anything in our inequality measures. Most equity measures satisfy this axiom (Cowell, 2011).

In the following sections we examine desirable properties that apply to equality of condition measures, impartiality measures and redistributive or meritocratic measures. No such properties apply to minimum standards measures, which consist simply of a binary variable – the standard is met or not – and the percentage of individuals who meet it.

2.5.1 Five desirable properties for equality of condition measures

1. Symmetry or anonymity

Symmetry or anonymity means that the measure is insensitive to any permutation of the variable of interest. If person A has three years of education and person B has six years of education, then swapping them around so A has six years and B has three years will not change the measure.

9 Reardon (2009) develops measures of segregation of *unordered* groups into *ordered* categories, for example of black and white students into different categories of educational attainment. These can also be considered measures of impartiality.

2. Perfect equality

An index of inequality of condition has a defined value which it takes when every individual has exactly the same level of education. Many indices are designed to range from 0 to 1, so 0 represents perfect equality and 1 represents some definition of perfect inequality.

3. The principle of transfer

Given a distribution of years of education, consider taking one year of education from an individual near the top of the distribution and adding one year of education to an individual near the bottom of the distribution. In general, inequality indicators should indicate that inequality has been reduced as a result of this transfer (Cowell, 2011).[10] An exception is if the individuals have now swapped positions so that the individual previously near the bottom is now nearer the top and *vice versa*.

Simple measures of dispersion, such as the range and inter-quartile range, have the virtue of being easy to calculate and understand, but the disadvantage is that they use only two points in the distribution and ignore everything that happens in-between. Transfers between any other individuals in the population will have no effect on the indicator. Even some measures that do take into account the education of every individual, such as the mean deviation, do not pass the principle of transfer test. Most of the indicators that remain commonly used, such as the Gini and variance, do pass this test.

Among those that conform to the principle of transfer, indicators give different weights to transfers at different points in the distribution. In particular, the Gini coefficient places a greater emphasis on transfers near the middle of a distribution than towards the top and bottom (Cowell, 2011). Some measures – such as the McLoone index, the standard deviation of logarithms, and (depending on the social welfare function chosen) the Atkinson index – place more weight on transfers affecting the lower end of the distribution (Berne and Stiefel, 1984), which may make them particularly relevant if we are trying to measure both inequality and the degree of deprivation of the most marginalised.

4. Scale and translation invariance

In income inequality, scale invariance is often taken to be an important characteristic of inequality measures. The "measured inequality of the slices of the cake should not depend on the size of the cake" (Cowell, 2011, p. 63). This means that multiplying everyone's income by the same amount would not change the inequality metric. The Gini, Theil and Atkinson measures all pass this test, but the variance and standard deviation, for example, do not.

The advantages of a scale-invariant inequality metric are clear: it separates inequality from the scale of the indicator and enables comparisons across contexts where the average value of the indicator may differ. However, it is not always clear that scale-invariant measures tell us what we are interested in knowing, even in income inequality (Ravallion, 2007; Hoy, 2015). For example, changes in the absolute gaps between rich and poor as an economy grows capture in more intuitive terms who is accumulating the extra income generated from growth (Hoy, 2015).

Similar arguments may be made in relation to the expansion of education systems. At the extreme, consider a society with two individuals, where one person never attends school and the other attends school for four years. Compare this to a society consisting of one person who never attends school, while the other person attends for eight years. Scale invariance means treating these two societies as equally unequal. The second society, one might argue, has a better-developed education system,

10 The explanation here borrows heavily from Cowell's (2011) textbook on income inequality, swapping income for education. The principle of transfer also has a "strong" form, which requires, in addition, that the amount of change in inequality due to a transfer depends only on the "distance" between individual ranks, and not on their position in the income distribution.

but in both cases one individual gets the maximum benefit while the other gets none. This argument is reasonable, but scale invariance also goes against a strong intuition that the second society is much more unequal, because the educational gap between the two individuals is simply much larger. In many policy discussions, this intuitive argument would be the most relevant one. A gap of eight years between the least and most educated individuals has much greater implications for the structure of society, different employment opportunities, and so on, than a gap of four years.

An alternative property that inequality measures may fulfil is translation invariance: if everyone receives one additional year of education, then the measure will not change. A society consisting of an individual with no education and an individual with four years of education is equally unequal to a society in which the individuals have one year and five years of education, respectively.

A measure that is "meaningful" – satisfying the symmetry and transfer properties, and continuous in any individual income – cannot be both scale-invariant and translation-invariant (Zheng, 1994). Thus, inequality measures tend to fall into one of two camps: "absolute", translation-invariant measures, and "relative", scale-invariant measures.

Measures that are not scale-invariant should usually be seen as being in the unit of the underlying indicator, such as dollars of expenditure per pupil or years of education. Additional care is therefore needed in making comparisons using non-scale-invariant measures. For example, comparisons in expenditure over time need to be adjusted for inflation; if they are not, then inequality may wrongly appear to be increasing (Berne and Stiefel, 1984). Similarly, comparisons across countries require expenditures to be converted into the same currency. With a scale-invariant inequality measure, such differences would not matter.

Many inequality measures have both absolute and relative versions. The difference often involves simply dividing or multiplying by the mean. For example, the standard deviation and variance are not scale-invariant, but dividing the standard deviation by the mean produces the coefficient of variation, which is scale-invariant. The Gini index, a relative measure, can be multiplied by the mean to produce the absolute Gini.

5. Decomposability

Decomposability implies that there should be a coherent relationship between inequality in the whole of a society and the inequality in sub-groups that make up that society. Notably, the Gini coefficient is *not* decomposable: it is possible for it to register increases in inequality in every sub-group at the same time as a decrease in inequality overall. Although mathematically attractive, this property is only really important if the intention is to decompose the indicator into sub-groups.

2.5.2 Five desirable properties of impartiality measures

In broad terms, the same set of principles applies to impartiality measures as to equality of condition measures. However, the exact meaning of these principles differs slightly, given that impartiality measures are inherently multivariate – that is, they depend on at least one "independent" circumstance variable, as well as at least one "dependent" educational variable – while equality of condition measures are univariate.

1. Symmetry or anonymity

As for equality of condition, we are not interested in permutations of individuals, provided the multivariate distribution remains the same. If A has six years of education and parental income of $200 and B has four years of education and parental income of $100, and we swap both education and parental income of the two individuals, then the impartiality measure will not change. If we were to swap their education levels

without swapping their income levels, however, then the measure might change.

2. Perfect equality

There should be a defined point which represents perfect impartiality, where there is no relationship between circumstances and education. Some impartiality measures can take negative values that tell us more about the nature of the relationship. The correlation coefficient, for example, ranges from -1 (perfect negative relationship between circumstances and education) to +1 (perfect positive relationship between circumstances and education), with 0 representing perfect impartiality.

3. The principle of transfer

Like the principle of transfer for equality of condition measures, a transfer principle may also apply for impartiality measures. A transfer of one year of education from a richer to a poorer child should always improve impartiality with respect to wealth.[11] Measures such as the regression slope fulfil this principle for linear regression, although not necessarily for more complicated regression specifications, such as those including a quadratic term.

4. Scale and translation invariance

Impartiality measures may be sensitive to changes in the scale of the circumstance variable(s) (e.g. parents' income) and in the education indicator. For example, a linear regression coefficient will increase if either the circumstance variable or the education indicator is increased by 10% for all individuals. A regression coefficient, also known as the sloped indicator of inequality (SII), is an *absolute* impartiality indicator and should be considered as having a unit that reflects this, such as "years of education per dollar of parental income". On the other hand, linear regression coefficients are translation invariant: they do not vary in response to an absolute change in either the independent or dependent variable. Adding an extra year of education to all children, or an extra dollar of income to all parents, will not change the regression slope.

Like univariate dispersion measures, impartiality measures can also be divided or multiplied by the mean to turn an absolute measure into a relative one, or *vice versa*. The relative inequality index (RII) is the slope divided by the mean of the dependent variable (e.g. years of education). It is scale-invariant with respect to changes in the educational outcome, but not with respect to the independent variable (e.g. parental income). An absolute version of the concentration index can also be calculated by multiplying it by the mean level of the outcome variable, such as years of education (Wagstaff et al., 1991). Elasticity – the estimated percentage change in education that would result from a 1% increase in parental income – is scale-invariant but not translation-invariant with respect to both education and parental income and is thus fully relative.

The correlation coefficient is a unit-less indicator, insensitive to changes in either scale or translation. As noted in Section 2.3.3, the correlation coefficient represents the strength of relationship rather than its magnitude and so is neither "absolute" nor "relative".

Impartiality often involves examining the extent to which inequalities of condition – such as the distribution of years of education – can be explained by circumstances. In such cases, it is also necessary to choose an underlying equality of condition measure. For example, Ferreira and Gignoux (2011) use the variance in PISA scores to measure inequality of condition, and the amount of that variance, explained by a set of circumstantial variables, as a measure of impartiality.

11 An exception to this is the unlikely case where the transfer reverses the whole relationship between wealth and education. For example, if education is positively correlated with wealth before the transfer, and negatively correlated with wealth after the transfer, then it might be that the transfer has worsened impartiality.

5. Decomposability

For impartiality measures, there are potentially two different forms of decomposition. As "inequality of opportunity" measures tend to measure the contribution of different circumstances to outcomes, it is useful to be able to decompose the total measure into the contribution of each of these circumstances. As an illustration, Ferreira and Gignoux (2011) present estimates of the contribution of each circumstance to mathematics scores in a number of countries; their measure is such that the measure of the contribution of each circumstance adds up to the total measure.

2.5.3 Desirable properties of redistributive and meritocratic measures

Meritocracy and redistributive measures are both forms of "appropriately unequal treatment of unequals" in the terminology of Berne and Stiefel (1984). As mentioned in Section 2.3.5, there are two distinct approaches for redistributive measures and the same can also be said for meritocratic measures. In the first, one has a particular extent of unequal treatment in mind, for example that poor students should get twice as much public expenditure as rich students, and can use this to weight the data. Univariate dispersion (equality of condition) measures can then be applied, with similar desirable properties to those discussed in Section 2.5.1. In the second approach, one does not have a particular expectation about how unequally to treat unequals but rather the desire to measure how unequally they are currently treated. In this case, measures such as the regression slope can be applied, and the desirable properties are similar to those used for impartiality *(see Section 2.5.2).*

2.6 SUMMARY

This chapter has presented a theoretical overview of the concepts of equity and inequality, and how these can be applied to the measurement of equity in learning and related educational variables. It discusses five key concepts for measuring equity in learning: meritocracy, minimum standards, impartiality, equality of condition and redistribution. But what an equity measure means in practice also depends on what one is measuring, and so we have also described some of the characteristics of educational variables that matter from an equity perspective. Finally, we set out a number of desirable properties that equity measures should meet, depending on how one aims to use them.

The following chapter illustrates how these concepts can be applied in practice, focusing on two particular equity concepts: impartiality and equality of condition.

REFERENCES

Atkinson, A. B. (1970). "On the measurement of inequality". *Journal of Economic Theory.* Vol. 2, No. 3, pp. 244-263. https://doi.org/10.1016/0022-0531(70)90039-6

Arrow, K. J. (1971). "A Utilitarian Approach to the Concept of Equality in Public Expenditures". *The Quarterly Journal of Economics*. Vol. 85, No. 3, pp. 409-415. https://doi.org/10.2307/1885930

Berne, R. and L. Stiefel (1984). *The Measurement of Equity in School Finance*. Baltimore, Maryland: Johns Hopkins University Press.

—— (1994). "Measuring equity at the school level: The finance perspective". *Educational Evaluation and Policy Analysis.* Vol. 16, No. 4.

Bransford, J. D., A. L. Brown and R. R. Cocking (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Cederman, L.-E., N. B. Weidmann and K. S. Gleditsch (2011). "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison". *The American Political Science Review*. Vol. 105, No. 3, pp. 478-495. https://doi.org/10.2307/41480853

Cingano, F. (2014). "Trends in Income Inequality and its Impact on Economic Growth". OECD Social, Employment and Migration Working Papers. Paris: OECD. http://www.oecd-ilibrary.org/content/workingpaper/5jxrjncwxv6j-en

Cowell, F. (2011). *Measuring Inequality* (Third Edition). Oxford: Oxford University Press.

Dumont, H., D. Istance and F. Benavides (2010). *The Nature of Learning: Using Research to Inspire Practice*. Paris: OECD Publishing.

Erreygers, G. and T. Van Ourti (2011). "Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice". *Journal of Health Economics*. Vol. 30, No. 4, pp. 685-694. https://doi.org/10.1016/j.jhealeco.2011.04.004

Ferreira, F. and J. Gignoux (2011). *The Measurement of Educational Inequality: Achievement and Opportunity.* IZA Discussion Paper No. 6161. Bonn: Institute for the Study of Labor (IZA). Retrieved from http://econpapers.repec.org/paper/izaizadps/dp6161.htm

Friesen, S., C. Saar, A. Park, C. Marcotte, T. Hampshire, B. Martin, B. Brown and J. Martin (2015). *Focus on Inquiry*. Calgary: Galileo Educational Network.

Gamboa, L. F. and F. D. Waltenberg (2015). "Measuring Inequality of Opportunity in Education by Combining Information on Coverage and Achievement in PISA". *Educational Assessment*. Vol. 20, No. 4, pp. 320-337. https://doi.org/10.1080/10627197.2015.1093926

Henriques, R. (2009). "Educational marginalization in Brazil". Background paper for the *EFA Global Monitoring Report 2010*. Paris: UNESCO.

Hinton, C. (2005). "A Report of the Learning Sciences and Brain Research". Third Lifelong Learning Network Meeting. Wako-shi, Saitama, Japan: OECD.

Hoy, C. (2015). *Leaving No One Behind: The Impact of Pro-Poor Growth.* Retrieved from https://www.odi.org/publications/9997-leave-no-one-behind-impact-pro-poor-growth

Hunt, F. (2008). *Dropping Out from School: A Cross Country Review of Literature*. CREATE Pathways to Access: Research Monograph no 16. Sussex: University of Sussex Centre for International Education. Retrieved 17 April 2017, from http://www.create-rpc.org/pdf_documents/PTA16.pdf

Illeris, K. (2007). *How We Learn: Learning and Non-learning in School and Beyond*. London: Routledge.

—— (2009). *Contemporary Theories of Learning: Learning theorists... in their own words*. London and New York: Routledge.

Jacob, W. J. and D. B. Holsinger (2008). "Inequality in education: a critical analysis". In D. B. Holsinger and W. J. Jacob (eds.), *Inequality in Education: Comparative and International Perspectives,* pp. 1-34. Hong Kong: Springer. Retrieved from http://sociologia.davidjustino.com/wp-content/uploads/2012/05/HOLSINGER2008_inequality_in_education.pdf

James, D. R. and K. E. Taeuber (1985). "Measures of segregation". In N. Tuma (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

Lassibille, G. and J.-P. Tan (2007). "Benefit Incidence Analysis in Education". *Journal of Education Finance*. Vol. 33, No. 2, pp. 170-182.

Massey, D. S. and N. A. Denton (1988). "The dimensions of residential segregation". *Social Forces*. Vol. 67, No. 2, pp. 281-315.

OECD (2012). *Equity and Quality in Education*. Paris: OECD. Retrieved from http://www.oecd-ilibrary.org/content/book/9789264130852-en

Østby, G. (2008). "Polarization, Horizontal Inequalities and Violent Civil Conflict". *Journal of Peace Research*. Vol. 45, No. 2, pp. 143-162. https://doi.org/10.1177/0022343307087169

Østby, G. and H. Urdal (2010). *Education and Civil Conflict: A Review of the Quantitative, Empirical Literature*. Paris: UNESCO. Retrieved from http://unesdoc.unesco.org/images/0019/001907/190777e.pdf

Paes de Barros, R., F. H. G. Ferreira, J. R. Molinas Vega and J. Saavedra Chanduvi (2008). *Measuring Inequality of Opportunities in Latin America and the Caribbean.* Washington, DC: The World Bank. Retrieved from http://documents.worldbank.org/curated/en/219971468045038979/Measuring-inequality-of-opportunities-in-Latin-America-and-the-Caribbean

Pfeffer, F. T. (2015). "Equality and Quality in Education. A Comparative Study of 19 Countries". *Social Science Research*, No. 51, pp. 350-368. https://doi.org/10.1016/j.ssresearch.2014.09.004

Phillips, A. (2004). "Defending equality of outcome". *Journal of Political Philosophy*. Vol. 12, No. 1, pp. 1-19.

Ravallion, M. (2007). "Inequality is Bad for the Poor". In S. Jenkins and J. Micklewright (eds.), *Inequality and poverty re-examined.* Oxford: Oxford University Press. Retrieved from https://pdfs.semanticscholar.org/339a/e202252a4bf09388320da2cd72a13da4808f.pdf

Rawls, J. (1971). *A Theory of Justice* (reissue edition). Cambridge, Mass.: Harvard University Press.

—— (2001) *Justice as Fairness: A Restatement.* Cambridge, Mass.: Harvard University Press.

Reardon, S. F. (2009). "Measures of ordinal segregation". In Y. Flückiger, S. F. Reardon and J. Silber (eds.), *Research on Economic Inequality.* Vol. 17, pp. 129-155. Bingley: Emerald Group Publishing Limited. Retrieved from http://www.emeraldinsight.com/doi/10.1108/S1049-2585%282009%290000017011

Roberts, F. S. (1979). *Measurement Theory: With Applications to Decisionmaking, Utility, and the Social Sciences* (Encyclopaedia of Mathematics and its Applications, Vol. 7). Cambridge: Cambridge University Press.

Roemer, J. E. (1998). *Equality of Opportunity*. Cambridge, Mass.: Harvard University Press.

—— (2002). "Equality of Opportunity: A Progress Report". *Social Choice and Welfare*. Vol. 19, No. 2, pp. 455-471.

Roemer, J. E. and A. Trannoy (2016). "Equality of Opportunity: Theory and Measurement". *Journal of Economic Literature*. Vol. 54, No. 4, pp. 1288-1332. https://doi.org/10.1257/jel.20151206

Sabates, R., K. Akyeampong, J. Westbrook and F. Hunt (2010). "School Drop out: Patterns, Causes, Changes and Policies". Background paper prepared for the *Education for All Global Monitoring Report 2011*. Paris: UNESCO. Retrieved from http://unesdoc.unesco.org/images/0019/001907/190771e.pdf

Sen, A. (1992). *Inequality Re-Examined.* Oxford: Clarendon Press.

—— (1999). *Development as Freedom*. Oxford: Oxford University Press.

—— (2002). "Response to Commentaries". *Studies in Comparative International Development.* Vol. 37, No. 2, pp. 78-86.

—— (2003). "Development as Capability Expansion". In S. Fukuda-Parr et al. (eds.), *Readings in Human Development*. New Delhi and New York: Oxford University Press.

Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*. Vol. 103, No. 2684, pp. 677-680. https://doi.org/10.1126/science.103.2684.677

Stewart, F. (2002). *Horizontal Inequalities: A Neglected Dimension of Development*. QEH Working Paper No. 81. Oxford: Queen Elizabeth House, University of Oxford. Retrieved from https://ideas.repec.org/p/qeh/qehwps/qehwps81.html

—— (2008). "Horizontal Inequalities and Conflict: An Introduction and some Hypotheses". In F. Stewart (ed.), *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. New York: Palgrave Macmillan.

Stewart, F. and A. Langer (2008). "Horizontal Inequalities: Explaining Persistence and Change". In *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230582729_4

Theil, H. (1967). *Economics and Information Theory*. Chicago: Rand McNally.

Toutkoushian, R. K. and R. S. Michael (2007). "An Alternative Approach to Measuring Horizontal and Vertical Equity in School Funding". *Journal of Education Finance*. Vol. 32, No. 4, pp. 395-421.

UNESCO Institute for Statistics (UIS) (2016). *Who Pays for What in Education? The Real Costs Revealed through National Education Accounts*. Montreal: UNESCO Institute for Statistics. Accessed August 2017 at http://uis.unesco.org/sites/default/files/documents/who-pays-for-what-in-education-national-revealed-through-accounts-2016-en_0.pdf

United Nations (1989). *The United Nations Convention on the Rights of the Child*. New York: United Nations.

—— (2017). General Assembly resolution 71/313, *Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development*, A/RES/71/313 (6 July 2017). Accessed 8 February 2018 at https://undocs.org/A/RES/71/313

UNESCO (2010). *Education for All Global Monitoring Report 2010: Reaching the Marginalized*. Paris: UNESCO.

Van Damme, D. (2014). "How closely is the distribution of skills related to countries' overall level of social inequality and economic prosperity?" EDU NAEC Paper Series. Paris: OECD.

Wagstaff, A., P. Paci and E. van Doorslaer (1991). "On the measurement of inequalities in health". *Social Science & Medicine*. Vol. 33, No. 5, pp. 545-557.

Walzer, M. (1984). *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.

Wilkinson, R. and K. Pickett (2010). *The Spirit Level: Why Equality is Better for Everyone*. London and New York: Penguin.

Zheng, B. (1994). "Can a Poverty Index Be Both Relative and Absolute?" *Econometrica*. Vol. 62, No. 6, pp. 1453-1458. https://doi.org/10.2307/2951756

3. Proposed operationalisation of equity measurement

BY CARINA OMOEVA, WAEL MOUSSA AND RACHEL HATCH
FHI 360 Education Policy and Data Center

3.1 INTRODUCTION

If equity can be conceptualised from a myriad of different perspectives, how, then, can it be measured? As this handbook proposes in Chapter 2, there are five different conceptual categories of equity measurement in education: meritocracy, minimum standards, impartiality, equality of condition and redistribution. This chapter focuses on the technical aspects of measurement and delves deeper into two key families of metrics: impartiality and equality of condition. The two groups are closely related, and in many cases are sufficient for reporting on key education indicators. As defined in Chapter 2 (*see Table 2.1*), the former denotes bivariate or multivariate associations between educational outcomes and sociodemographic characteristics, while the latter reflects univariate inequality metrics.

Chapter 2 also emphasises that there is an important distinction between the concepts of equality and equity, with the first a mere state of being equal with respect to a given input, characteristic or outcome, while the latter considers the social and economic context and introduces the concept of fairness to the concept of equality. Thus, *equity* can be seen as a pathway towards greater *equality* in education outcomes; and whereas perfect equality among all individuals cannot be achieved (nor is always desirable), greater equity would mean that the distribution of outcomes is more independent of the socioeconomic or cultural characteristics of students or their schools.

In this chapter, we guide the reader through a common process for identifying relevant equity considerations and provide the basic technical and operational framing for some of the most common metrics for equity that can be applied towards education indicators. The concept of *inequality* is used here to denote the simple condition of a lack of mathematical equality, which should further lead the reader to explore whether this inequality has social and economic dimensions. If it does, this would indicate the presence of inequity. The reader is introduced to visualisation and measurement techniques that represent equality of condition (univariate measures of inequality) and impartiality (bi- or multivariate measures of inequality). We begin with an initial, dimension-agnostic visual process for gauging inequality, and proceed using each of the highlighted approaches through several examples drawn from country- and programme-level data.

In each case we address the requirements needed for using the underlying data to measure (in)equality of condition and impartiality, as well as the appropriateness of the given measure – its advantages and disadvantages – for generating insights into the magnitude and nature of the inequality. The chapter continues with a discussion of key metrics of inequality, such as measures of disparity, dispersion and more complex indices like Gini and Theil that require continuous and cumulative information. It offers an overview of the current equity in education landscape and makes some recommendations for improving data availability, as

well as listing a simple sequence of steps to follow for analysing equity.

3.2 VISUAL REPRESENTATIONS OF EQUALITY OF CONDITION

Any analysis of equity in education should begin with a simple visual analysis of the distribution of the indicator in question. This is done by examining the extent to which it is equal or unequal across observations within a given dataset. While perfect equality can never be reasonably expected, the nature and shape of distributions and the magnitude of inequality contained within them can present an initial indication of whether more analysis is necessary. The measurement of inequality can be characterised as a measure of the degree of homogeneity or heterogeneity within a certain population and between groups in terms of educational output. Throughout this chapter, we define educational output in terms of learning outcomes (assessments, test scores), which are continuous outcomes, and the completion of specific educational milestones such as graduation, progression, drop-out, etc., which are dichotomous outcomes. The higher the degree of heterogeneity, the higher the degree of inequality between students or between groups. Further analysis can also help gauge the equality of condition or impartiality. Among their many advantages, graphical representations are also useful in identifying outlier behaviour that may not always be captured using numerical statistical analyses.

In this sub-section, we provide the reader with a number of graphical analytic tools as initial gauges of univariate inequality (equality of condition). We describe three well-known methods of representing inequality and discuss their advantages and disadvantages. In addition, each sub-section will provide theoretical illustrations as well as empirical examples from available learning outcomes data. Visual representations of impartiality are shown throughout the section that uses bivariate methods for representing inequality at the group level.

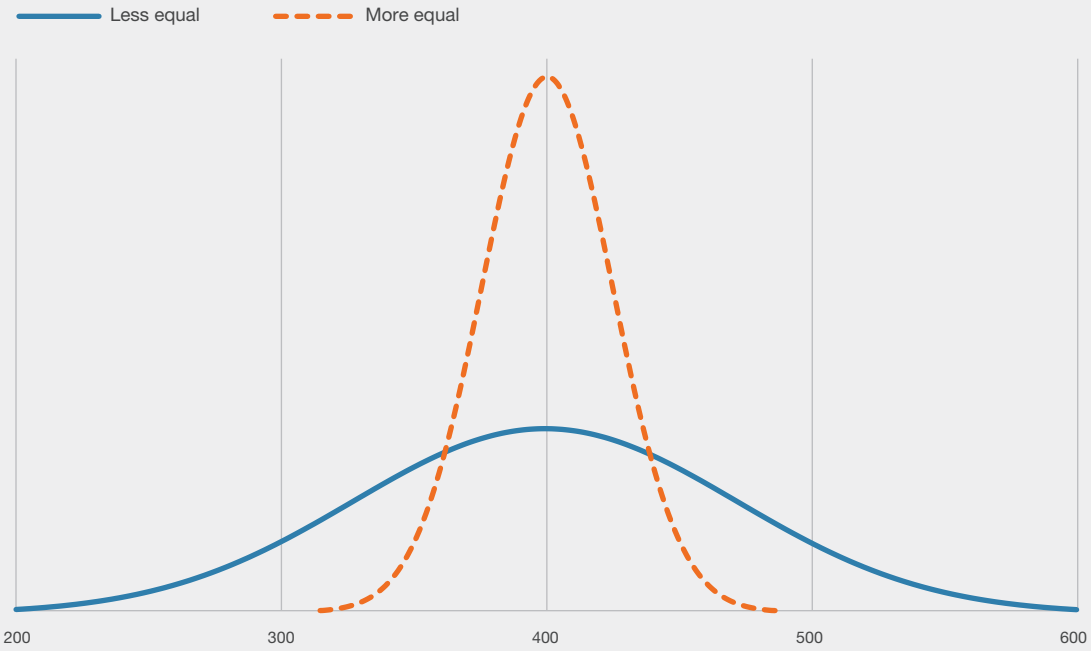
Finally, because univariate representations of binary educational outcomes at the individual level are trivial, they are not included in this chapter.

3.2.1 The histogram or the probability density function

We begin our review with what is perhaps the simplest way of representing dispersion and inequality: the histogram. Histograms are a graphical representation of a probability density function (PDF) for continuous outcomes and a probability mass function (PMF) for discrete outcomes. It is a visual representation of the distribution of a given population across all possible values of a given outcome. The use and applicability of a histogram or a representation of a PDF would therefore depend on the type of learning or educational outcomes data available and applicable only to univariate data analyses. In this case, a histogram is only applicable to continuous individual-level data as a histogram will not discriminate along sub-group membership. This is because it can only show the proportion of the entire population who have realised a certain level of an outcome, which makes it inapplicable to group-aggregated data or bivariate associations. Although a histogram on its own may not inform the absolute level of inequality that exists within a given context, a histogram alone serves the purpose of illustrating the degree of dispersion that exists on, say, a reading test score, and can be compared to a "theoretical ideal" distribution or an existing empirical benchmark distribution.

Figure 3.1 illustrates the distribution of a hypothetical learning assessment score that ranges between 200 and 600. In this figure, we plot two distinct distributions that represent different levels of inequality. In addition, we include a solid vertical line to represent perfect equality that serves as a benchmark to illustrate the gap between perfect equality and the actual empirical distribution. We note that, in the case of assessment results, perfect equality would necessarily mean that all test-takers received the

Figure 3.1 Hypothetical test score distributions with varying levels of inequality



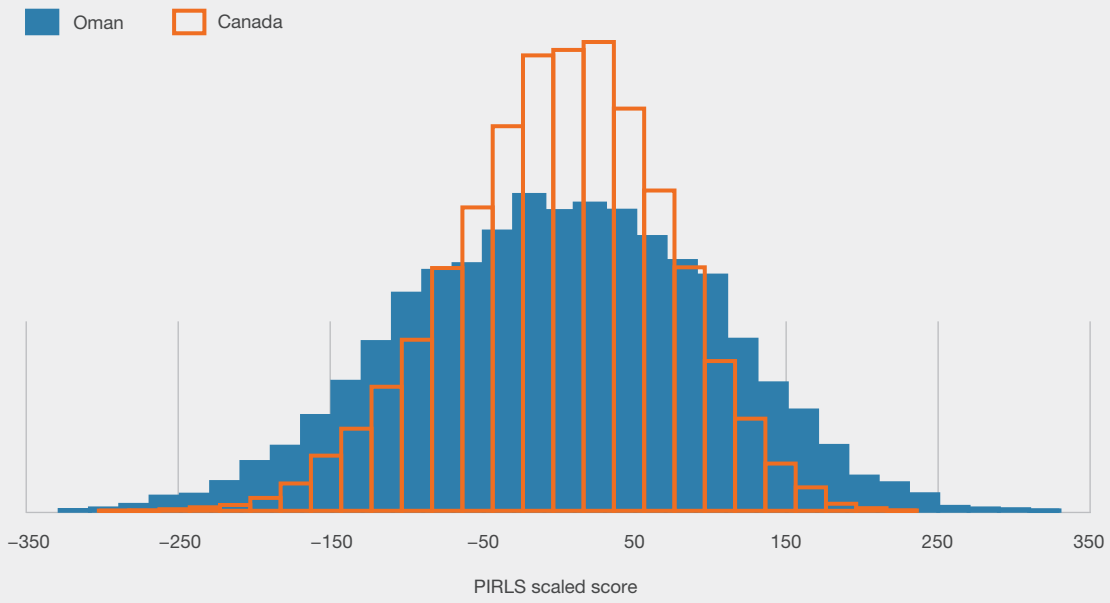
exact same score – in other words, 100% of the population achieve the same outcome. The figure shows the distribution represented by the solid curved line as being "less equal" as it provides a clear indication that the dispersion in observed test scores is much higher than those of the "more equal" dashed curved line. Therefore, as the distribution of scores tends toward the vertical line (narrower distribution), the condition of the population becomes more equal.

Using empirical data from the Progress in International Reading Literacy Study (PIRLS), we provide a similar application using a histogram to illustrate the degree of inequality, relative to a high-standard benchmark, as well as relative to perfect equality *(see **Figure 3.2**)*. For the purposes of analysing and isolating inequality, we rescale the PIRLS test scores to create mean-centred outcomes. This enables us to make a direct comparison of the degree of inequality between two populations, rather than comparing differences in mean achievement between two distinct groups of students. The PIRLS scores now have a mean of zero, but the variance around the mean remains unchanged.

From Figure 3.2 we can observe that the PIRLS reading assessments generate a normal distribution of scores around the mean for each country. In this case we examine PIRLS scores from Canada and Oman and compare their levels of score dispersion. The histogram shows that the overall distribution for Canada is narrower and has a higher peak around the mean, while Oman has a wider spread and a lower peak around the mean. This finding highlights the usefulness of a simple graphical method, such as the histogram, to identify distributions that are less or more equal relative to a given standard or benchmark.

Despite its usefulness for visualising the degree of inequality present within a population, a histogram

Figure 3.2 Empirical distributions of PIRLS test scores for Canada and Oman



Note: PIRLS test scores are mean-centred.
Source: Progress in International Reading Literacy Study (PIRLS), 2011.

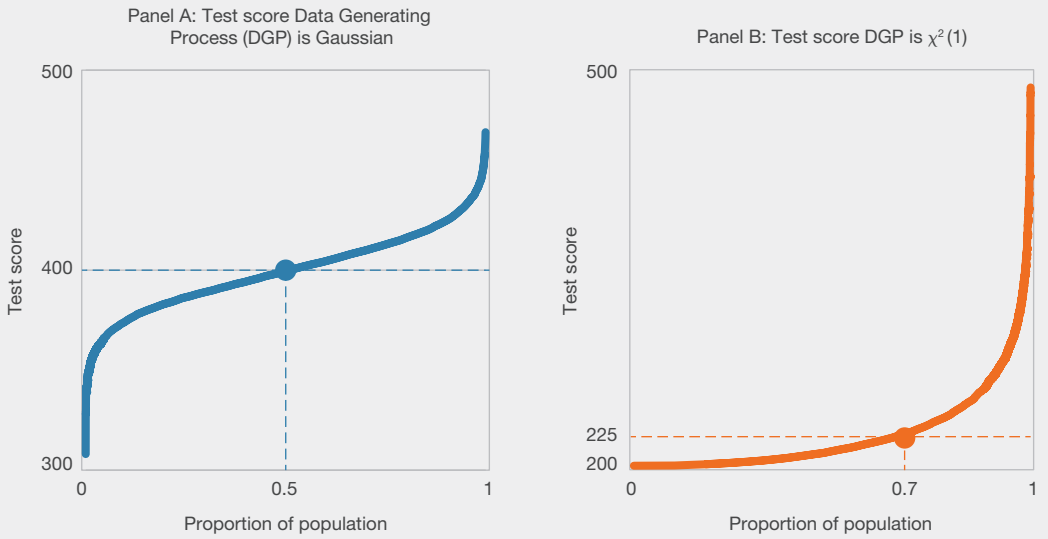
is not sufficient as a stand-alone tool for probing the nature or sources of the inequality. These drawbacks become apparent when a benchmark "distribution" is not present, because a histogram only identifies non-perfect equality, providing a relative sense of inequality when the distribution of a population of interest is coupled with that of a benchmark or standard distribution. Further, like all graphical analyses, a histogram does not provide a summary statistic to capture the degree of inequality of condition on an absolute scale. Nevertheless, it is a useful and informative tool as a first step in any data analytic exercise, especially in the context of visualising equality of condition of any given continuous educational outcome.

3.2.2 The cumulative distribution

Another well-known visualisation of equality of condition is the cumulative distribution function (CDF). The CDF is defined as the cumulative proportion, or frequency, of a population that has attained a certain level of educational output or less. It provides a direct method of measuring the proportion or number of students who have achieved up to a certain score and plots that proportion or number for every possible value of the assessment. For example, the CDF can be used to determine what percentage of the population scored below the lowest score on a standardised test, such as 400 on the SAT higher education entrance examination in the United States.[12]

12 The SAT is a standardised test used for post-secondary/university admissions in the United States. Students typically sit for the SAT assessment during their third year of high school, i.e. in the 11th grade.

Figure 3.3 Hypothetical cumulative distributions under different distributional assumptions



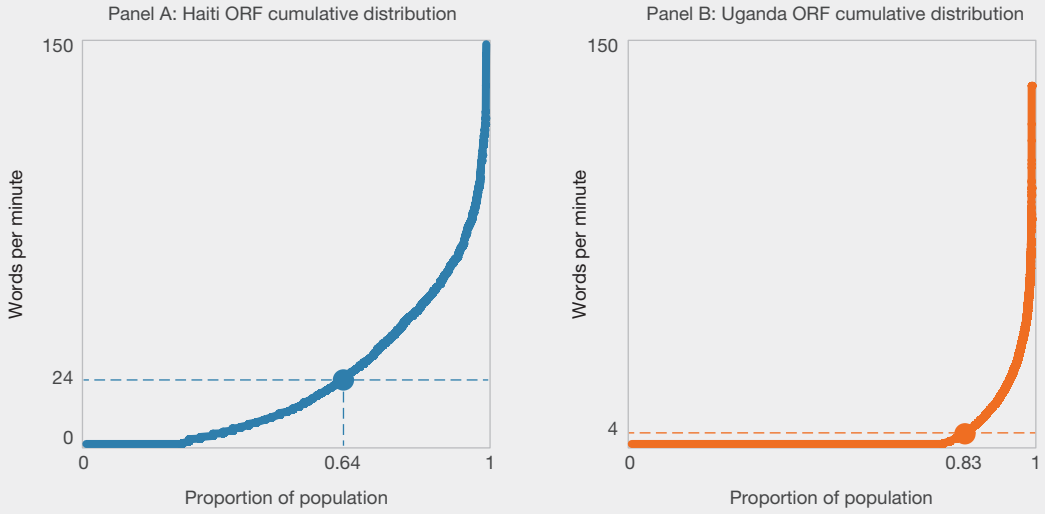
Note: Data Generating Process (DGP) means that we use different distributional assumptions for how the data are generated for the examples used throughout this chapter.

Figure 3.3 plots two cumulative distributions of a hypothetical learning assessment with possible scores ranging between 200 and 500, under two different distributions. These are not based on real data but rather are generated under different distributional assumptions for illustration. The figure in Panel A displays the CDF of a test score distribution that follows a Gaussian distribution function, with a mean of 400 and a standard deviation of 20. We can see in this case that 50% of the population have a score less than or equal to the mean score. However, it is unclear whether this indicates a high or low degree of inequality in terms of test score performance. Panel B displays the CDF of a test score distribution that is positively skewed, meaning that the majority of students have low test scores while a small number of students have much higher test scores.[13] The CDF of the skewed distribution shows that the mean is much closer to the lowest obtainable score and that 70% of the population have a score lower than the mean. This distribution clearly shows that, in terms of test score performance, the condition of the students is relatively unequal and illustrates the concept of Pen's parade.[14]

In addition, cumulative distribution graphs provide a simple representation of perfect equality of condition: in the case of learning assessment scores, perfect equality refers to a distribution where all students received an identical score and, as such, is represented by any straight line that is parallel to the horizontal axis. In other words, 100% of the population achieve the same outcome. This line can serve as a benchmark to compare the empirical distribution to a hypothetical perfect equality line. The larger the distances between each point on the CDF and the perfect equality line, the greater the degree of inequality in learning outcomes.

13 The test score distribution in Panel B is generated following a distribution with a mean of 225 and one degree of freedom.
14 Pen's Parade is a concept created by Jan Pen (1971) to describe the income distribution in an economy.

Figure 3.4 Empirical cumulative distributions of EGRA oral reading fluency (ORF) results for Haiti and Uganda



Source: Progress in International Reading Literacy Study (PIRLS), 2011.

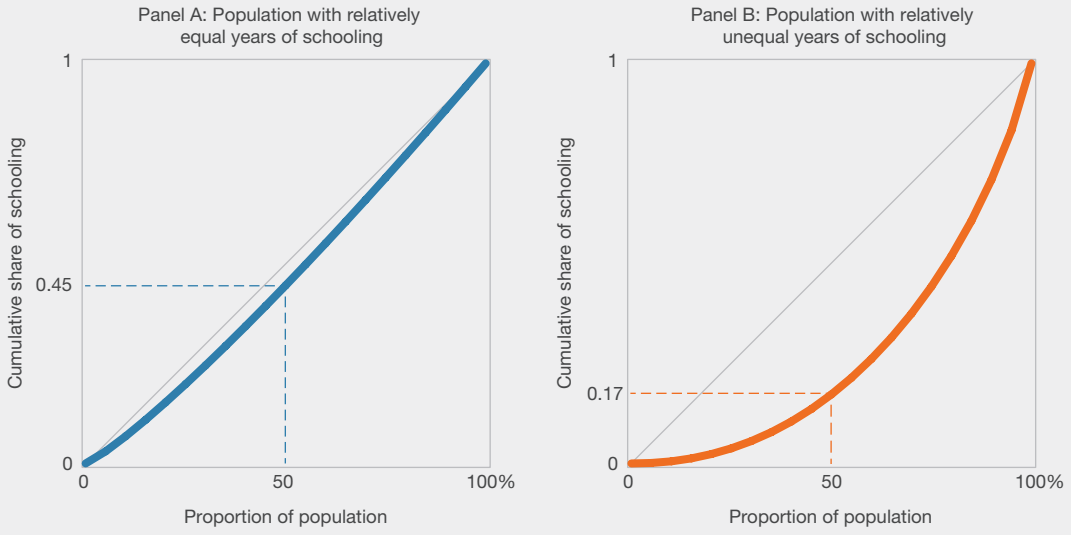
In the following example, we replicate the CDF analysis using data from Early Grade Reading Assessment (EGRA) administrations in Haiti and Uganda. As in Figure 3.3, **Figure 3.4** includes two panels representing the CDF. We chose EGRA oral reading fluency results to illustrate the CDF's usefulness because EGRA results typically follow positively-skewed distributions with a large portion of students receiving zero scores. Panel A of Figure 3.4 shows the CDF for Haiti where about 23% of the test-takers received a score of zero. At the same time, 64% of all test-takers are able to read 24 words correctly in a minute or less, which is also the mean oral reading fluency score in Haiti. In general, when more than one-half of the sample population's outcomes are lower than the mean, this indicates a relatively high degree of inequality. Thus, the higher the proportion of the population that is below the mean, the higher the degree of inequality. As an extreme example, the highest possible inequality would be a case where nearly all students receive the lowest possible score and only one student has a non-zero result, at which point almost 100% of the population have a score lower than the mean.

In Uganda, 83% of all test-takers are able to read correctly four words per minute or less, which is also the mean ORF score for the entire population. This result suggests two straightforward conclusions. First, Uganda exhibits a high degree of inequality with the vast majority of the student population scoring below average. And second, Uganda exhibits a higher degree of inequality than Haiti based on these results because a higher proportion of its population scores below average.

3.2.3 The Lorenz curve

One of the most popular visualisation tools when analysing equality of condition is the Lorenz curve, a graphic that follows naturally from the cumulative distribution function. The main difference between the CDF and the Lorenz curve is that the y-axis represents each proportion of the population's share

Figure 3.5 Hypothetical Lorenz curves illustrating two populations with different distributions of years of schooling



of the grand total of a given educational outcome. In other words, the Lorenz curve represents the share of the overall "pie" that belongs to each proportion of the population. Similar to the PDF and CDF representations of learning outcome distributions, the Lorenz curve allows researchers to easily identify the perfect equality line, which is represented by the 45-degree line from the origin and corresponds to a population where every individual owns an equal share of the "pie". Conversely, perfect inequality would be a right-angled line where one individual owns the entire "pie" and is thus represented by the curve farthest from the perfect equality line.

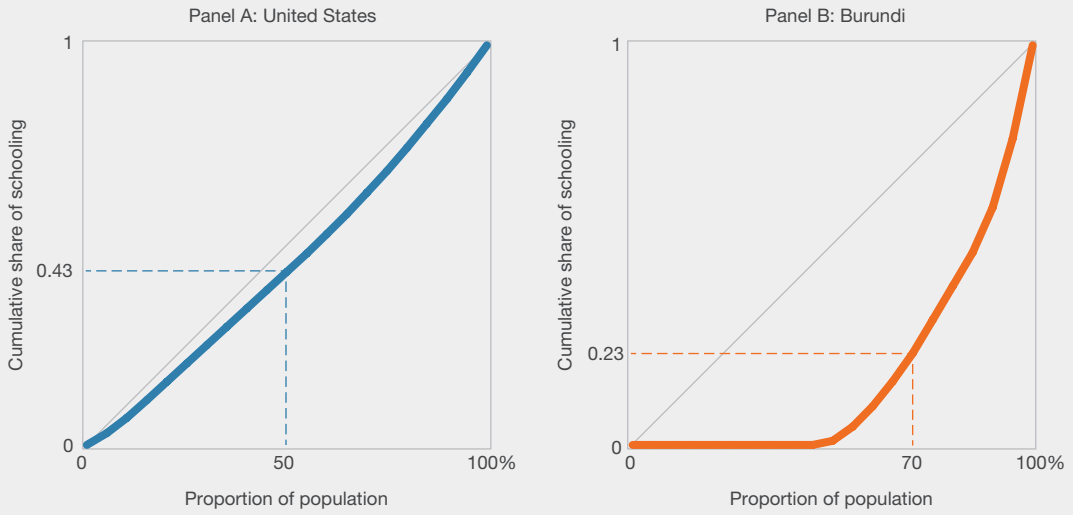
The popularity of the Lorenz curve stems from its straightforward interpretability. The closer it gets to the 45-degree line, the greater the equality and vice versa. In addition, the Lorenz curve function is the core component when calculating the popular Gini coefficient.[15] In this sub-section, we use completed years of schooling as a continuous random variable taking on values between 0 and 20 years to illustrate the Lorenz curve. **Figure 3.5** plots the Lorenz curve from two hypothetical distributions of years of schooling, where the first is based on a population with nearly everyone having the same years of schooling and the second is based on the converse situation.

Panel A plots the Lorenz curve for the population with a fairly equitable schooling distribution, where 50% of the population owns about 45% of the total stock of schooling. Panel B plots a population with a more unequal distribution of schooling, as 50% of the population owns merely 17% of the total stock of schooling. The Lorenz curve in Panel B is far from the perfect equality line, indicating the greater degree of inequality in years of schooling among its population.

In **Figure 3.6** we use empirical data on years of schooling from Burundi and the United States to plot the Lorenz curves for the two countries. The United States has a more equitable distribution of years of schooling among its population as the corresponding

15 The Gini coefficient will be described in detail in the next section of this chapter.

Figure 3.6 Lorenz curves for years of schooling in the United States and Burundi



Sources: United States Population Census 2010 and Burundi Demographic and Health Survey 2010.

Lorenz curve is close to the perfect equality line (45-degree line). On the other hand, Burundi exhibits a high degree of inequality as 70% of its population own only about 23% of the total stock of schooling in the country. Panel B in Figure 3.6 also shows that in Burundi almost 50% of the population own 0% of the schooling "pie".

3.2.4 Using visualisations of equality of condition

Visual representations of inequality provide an overview of the distribution of a given educational outcome and offer a general gauge of dispersion within a distribution. However, visual distributions of inequality do not provide a summary indicator of the magnitude of inequality and therefore are merely the first step in analysing equity. In the following sections, we delve into numerical measurements of equality of condition and impartiality from the perspective of univariate educational outcomes at the level of individuals and from the perspective of bivariate associations of educational outcomes and sociodemographic groupings at the aggregate/group level.

3.3 MEASURING INEQUALITY: A CATALOGUE OF COMMON METRICS

In this section, we discuss summary measures of inequality that describe, in various forms, existing disparities and gaps in educational outcomes relevant to different forms of research. The following family of metrics allows researchers to gauge the extent of disparities in educational outcomes across the population studied, without knowledge of the characteristics of the population itself. This section continues the discussion from the previous section on "plotting inequality" as a way to avoid relying solely on inadequate visual tools to identify and determine the degree of inequality within a population, or between groups of a population, in an all-encompassing summary form. The measures described lend themselves to analyses of both individual and aggregate-level data. **Table 3.1** lists these metrics in ascending order from simple to increasingly complex.

Table 3.1 Selected measures of inequality

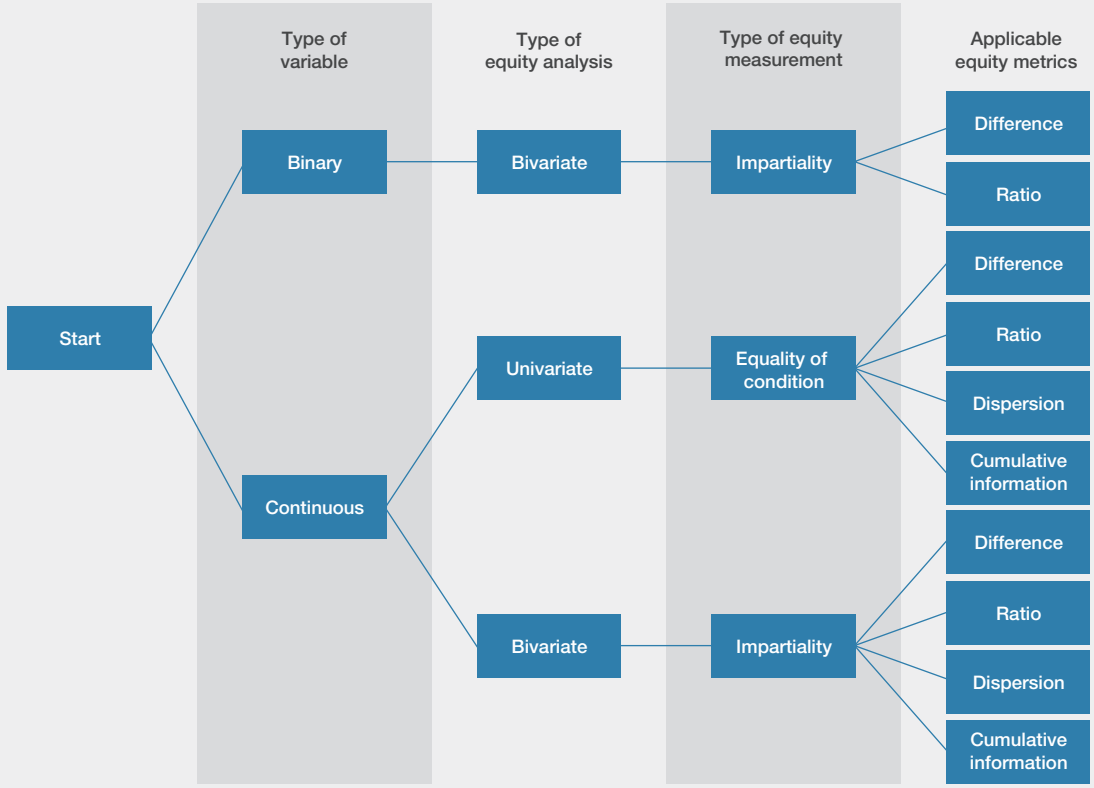
Concept	Family of metrics	Equality of condition measures	Description and remarks
Impartiality	Difference, gap	Range	Difference between highest and lowest indicator, across individuals, schools, districts, etc.
		Restricted range	Difference in the indicator at specific percentiles in the distribution, e.g. interquartile range is the difference between the 75th and 25th percentiles
	Ratio	Parity indices	The gender parity index is the ratio of female to male values of a given educational outcome (gross enrolment ratio, literacy rate, etc.)
		Palma ratio	Ratio of the share of education of the top 10% of the distribution to that of the bottom 40%. Variants use other percentiles. Considered policy-relevant for income but not yet applied in education.
Equality of condition	Dispersion	Variance or standard deviation	Average squared deviation (difference from the mean) in the indicator
		Coefficient of variation	Standard deviation divided by the mean
		Mean absolute deviation	Average absolute deviation (difference from the mean) in the indicator
	Cumulative information	Atkinson index	Cumulative inequality metric that determines the source of the inequality to be from the high end or the low end of the distribution
		McLoone index	Cumulative sum of indicator values for individuals below the median divided by the cumulative sum of the indicator for the same individuals as if they were at the median
		Gini coefficient	Relationship between the actual distribution and perfect equality in the outcome
		Theil index	Generalized entropy measure

Note: These measures are also presented in Tables 2.2 and 2.3 in Chapter 2.
Source: Authors' analysis.

Figure 3.7 puts Table 3.1 into context by illustrating the decisionmaking process to determine which of these equity metrics to use. **Table 3.2** catalogues common educational outputs and indicators pertinent to international development and their corresponding applicability to the different families of inequality metrics. The table details whether each indicator lends itself to measures of impartiality, equality of condition or both. However, when coupled with a socioeconomic/demographic characteristic or grouping, it enables an analysis of impartiality between groups of individuals.

This section discusses the data requirements, calculation methods, interpretation of the summary statistics, advantages and disadvantages, and empirical examples for each of the listed measures. The list covers only the most common metrics used

Figure 3.7 Decision tree for equity analysis and applicable metrics



Source: Authors' analysis.

in international education to analyse equity but should offer a way to compare the benefits and shortcomings of each. Last, this section is divided into two streams: measurements of equality of condition (univariate individual-level analysis) and measurements of impartiality (bivariate group/aggregate).

3.3.1 Equality of condition

3.3.1.1 Differences

One of the most intuitive and simple measures of inequality is the analysis of differences or gaps. Since this section discusses equality of condition at the individual level, calculation of gaps in educational outcomes requires data at the person level and a basic working knowledge of statistics. The mathematical definition of a difference, or a gap, is the distance in a straight line between two points. The formula for the difference, or a gap, in educational outcomes is as follows:

$$D = Y_i - Y_j \qquad [1]$$

where D represents the difference between the outcome realised by person i and person j. One of the main advantages of the difference measure is its flexibility since differences can be calculated between any two points. Obviously, this type of measurement is most useful when educational outcomes of interest are continuous in nature, since binary outcomes will result in a correspondingly trivial difference.

A common approach to summarising the degree of equality of condition in each educational outcome is to calculate the range. The range is the difference

Table 3.2 Common education indicators and their applicability to inequality measurement

Indicators	Binary, categorical	Continuous	Difference	Ratio	Dispersion	Cumulative information
Access and participation						
Gross enrolment ratio (GER)	X	X*	I	I		
Net enrolment rate (NER)	X	X*	I	I		
% of children ever accessing school	X	X*	I	I		
Completion						
Repetition rate	X	X*	I	I		
Drop-out rate	X	X*	I	I		
Survival rate	X	X*	I	I		
Completion rate	X	X*	I	I		
Transition rate to next education level	X	X*	I	I		
Learning						
% of students who are literate	X	X*	I	I		
% of students literate in mathematics	X	X*	I	I		
% of students passing national exams	X	X*	I	I		
% of students achieving minimum proficiency	X	X*	I	I		
Mean assessment score		X	I, EC	I, EC	I, EC	I, EC
Attainment						
Attainment (e.g. highest level/degree attained)	X		I	I		
Years of schooling		X	I, EC	I, EC	I, EC	I, EC
Resources						
Pupil-teacher ratio		X*	I, EC*	I, EC*	I, EC*	I, EC*
Government education spending		X	I, EC*	I, EC*	I, EC*	I, EC*
Household education spending		X	I, EC*	I, EC*	I, EC*	I, EC*

Notes: "I" refers to "impartiality" and "EC" refers to "equality of condition".
* denotes group-level data, thus "EC*" denotes equality of condition when the unit of observation is not the individual, e.g. school, municipality, region or district level.
X refers to variables being used as binary or categorical.
X* refers to binary variables that have been aggregated and become continuous variables as a result.
Source: Authors' analysis.

between the highest and lowest values obtained from a given population. The larger the magnitude of the range the higher the level of inequality. However, this method of summarising inequality can be misleading in the presence of outliers. In data from the PIRLS 2011 administration in the United Arab Emirates (UAE), the range between the highest and lowest scores is 745 points $(D = Y_{max} - Y_{min} = 801 - 56)$. This range or gap indicates a very high degree of inequality in reading proficiency among students in the UAE. However, upon inspection of the test score data, we observe that the majority (90%) of UAE students' scores fall between 274 and 605, both of which are substantially different from the respective minimum and maximum scores.

To circumvent the issue of outliers, one can maintain the usefulness of differences by calculating a restricted range such as the inter-quartile range (IQR). The IQR measures the difference between the

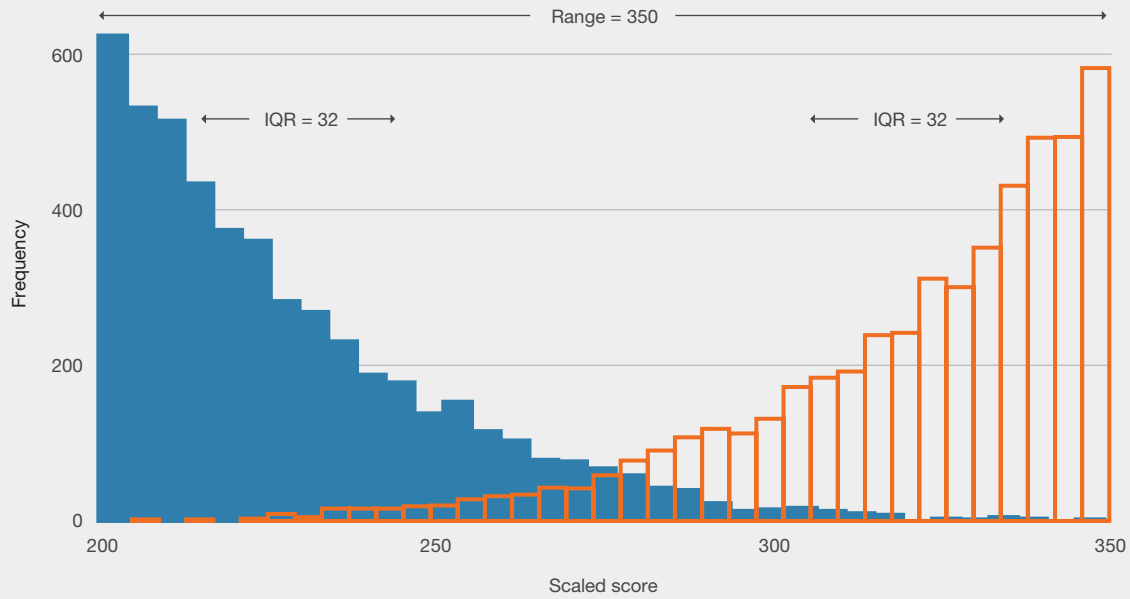
values of a variable at the 75th and 25th percentiles—in other words, the range of the middle 50% of the distribution. Another application of differences includes measuring the gap between individual outcomes and specific benchmarks, although the benchmarks are usually context-specific and may not be comparable across contexts.

It is important to note that, even with the relative ease of using the distance metric, there are two additional drawbacks in computing differences or restricted ranges. The first is scalability, where the scale of the educational outcome of interest plays a role in determining the magnitude of the calculated gaps. Thus, the relative magnitude of the gap can be ambiguous. For instance, an IQR of 150 scaled points on the PIRLS assessment does not inform on the magnitude of this gap between a student in the 75th percentile and a student in the 25th percentile if the significance of the 150 points cannot be determined without knowledge of the assessment's scaled score distribution and the relative importance of each assessment item. Given the non-linear nature of some assessment scales, it is also difficult to ascertain whether 150 points is a small or large difference.

The second, and more apparent, drawback is that measuring differences relies on two points of data only, while most of the information provided by the data is ignored. In **Figure 3.8**, we illustrate two distinct distributions of simulated test score data and show that the range and IQR are identical, whereas different measures of inequality would conclude otherwise. The figure plots the histogram of two contrasting distributions of simulated test score data, where the first distribution (shaded bars) is skewed towards the lowest possible score, which has a higher level of inequality (refer to the previous CDF sub-section for this determination); and the second distribution (blank bars) is skewed towards the highest possible score and is considered more "equal" based on the CDF.

Under both distributions, we calculate the range to be 150 $(D = Y_{max} - Y_{min} = 350 - 200)$, which

Figure 3.8 Hypothetical distributions of simulated test scores under different distributional assumptions



is due to the presence of at least one individual at the highest obtainable score and one at the lowest obtainable score, despite differences in the distribution skewness. Further, even when trying to circumvent the issue of extreme values in the distribution and calculating the IQR, we find the same result. The computed IQR for both cases is 32 points. Relying on the range or IQR would result in the conclusion that both populations exhibit an identical level of inequality, although we have shown previously that using the CDF leads to a different conclusion.

3.3.1.2 Ratios

The ratio is a commonly-used measure in education, especially in gender studies. Ratios are simply the result of dividing the value of an indicator for one person (or group) by that of another. This measure is similar to the difference metric in its simplicity and ease of interpretation, yet provides an advantage in that the relative magnitude of the disparity between two persons (or groups) can be identified. The formula for the ratio metric is as follows:

$$R = \frac{Y_i}{Y_j} \qquad [2]$$

where R represents the ratio of the value of a variable Y for individual i over that for individual j, and can be expressed in percentage terms or as a factor. Thus, the issues of scalability become less relevant when using ratios to determine the degree of inequality within a population.

With individual-level data from a Living Standards Measurement Study conducted in Nigeria in 2013, we calculate the ratio by comparing the years of schooling completed by an individual in the 90th percentile to the years completed by a person in the 10th percentile. The data show that a person in the top decile of attainment completed 17 years of schooling and a person in the bottom decile completed 6 years. The resulting ratio is 2.83, which means that a person in the top decile has 2.83 times the years of schooling completed relative to a person in the bottom decile. One may interpret this disparity between the top and bottom deciles in years of schooling to be 283%. Since the unit of measurement is in percentage terms, the scale of the indicator of interest is no longer relevant. Another common ratio-based metric is the Palma ratio, which is defined as the ratio of the education share of the top 10% to that of the bottom 40%.

The use of ratios to quantify inequality suffers from some of the same drawbacks found in gap calculations. Ratios, for example, also require only two observations for calculation, which means that all other information is ignored. Ratios can also be susceptible to outliers depending on the point of reference being used. For instance, maximum/minimum ratios will provide a distorted image of inequality in the presence of isolated extreme values. Caution is needed when calculating ratios for the purposes of representing inequality within a population, as the scale and nature of an indicator can contribute to misrepresenting equality of condition. It should be noted that ratios are not very useful when analysing assessment scores that do not follow a linear scale, e.g. scaled scores. Since common assessment practices weight items differently, the scoring mechanism is no longer a linear step function. Ratios are a linear construct and thus only meaningful if the indicator itself follows a linear scale.

3.3.1.3 Dispersion

Variance and standard deviation

A natural progression from fairly simple methods of measuring equality of condition leads to variance and standard deviation. The immediate advantage of variance over simple differences or ratios is that this method of measurement uses every data point available, thereby getting the most out of the information provided. Variance is defined as the average squared distance between the educational outcome of each individual in a population and the

population mean. Thus, the scale of the metric is in line with the scale of an education variable and includes information from every member of a population. Formally, variance and standard deviation are computed as follows:

$$V(Y) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 \qquad [3a]$$

$$SD(Y) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \qquad [3b]$$

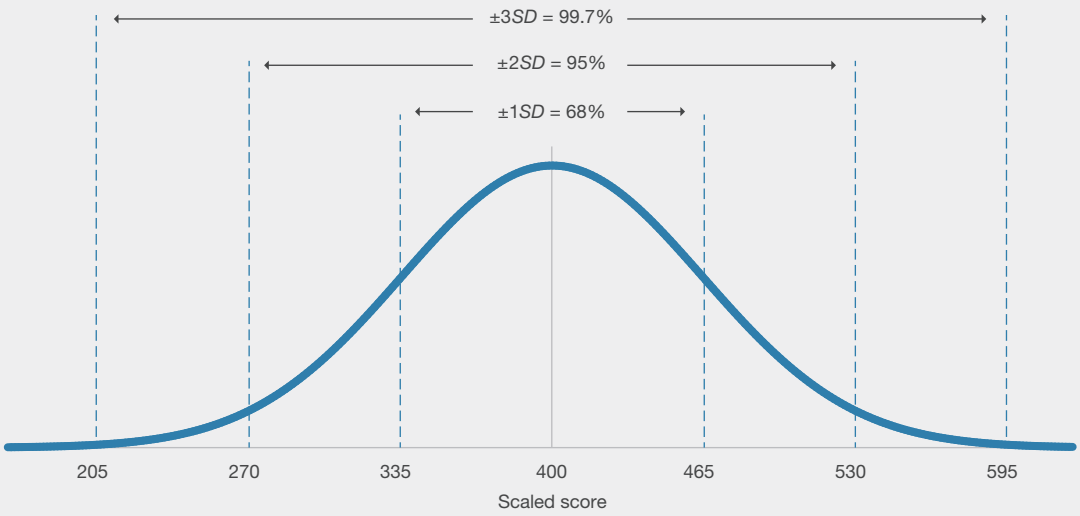
and

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad [3c]$$

where n represents the population size, Y_i is the value of an education variable for individual i, and $\overline{Y}$ is the mean of that education variable for the population. Variance, or the standard deviation metric, has several advantages. The measure is widely used and intuitive when coupled with knowledge of Chebychev's rule, where proportions of a normally distributed random variable can be divided into standard deviations. **Figure 3.9** illustrates this rule.

One standard deviation above and below the mean includes 68% of the total distribution, with one standard deviation representing 34% of the population. The magnitude of the standard deviation thus reveals the degree of clustering around the mean of a given distribution. For the concepts of variance and standard deviation to be most informative and meaningful, certain distributional requirements have to be met. One requirement is that the values of an education indicator must follow a normal (Gaussian) distribution; otherwise, a standard deviation can refer to varying levels of concentration depending on the shape of the distribution. One drawback to the use of variance as a measure of dispersion or inequality is its sensitivity to scaling. Suppose a researcher is interested in studying the distribution of educational attainment, measured in years of schooling, and wants to measure the variance of this distribution. If we were to double the number of years of schooling completed for the entire population, the variance would in fact quadruple. To illustrate, let Y represent

Figure 3.9 Gaussian distribution in relation to Chebychev's rule



the years of schooling variable and Y^* represent years of schooling after doubling Y for all individuals in the population, where $Y^* = 2Y$. The variance of Y is represented in equation [3a] and the variance of Y^* is as follows:

$$V(Y^*) = \frac{1}{n}\sum_{i=1}^{n}(Y^*_i - \overline{Y}^*)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(2Y_i - 2\overline{Y})^2 \qquad [4]$$

$$= \frac{4}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

$$= 4V(Y)$$

We can see that even when increasing years of schooling for all members of the population by an equal amount, the variance indicates that the degree of inequality has increased to 400% of the original variance. Using a standard deviation instead would indicate that inequality has risen twofold. It is therefore important to use caution with this metric, as it may not be appropriate for assessing changes in the magnitude of inequality over time.

Coefficient of variation

To avoid the drawbacks of variance and standard deviation, the coefficient of variation can be used to standardise the variance, thus making this measure of inequality immune to scaling issues and lack of relative magnitude. The coefficient of variation is defined as the standard deviation of an education indicator divided by its mean. As a result, the variable of interest is always a function of the mean, and the actual scale of the variable is no longer relevant. The coefficient of variation can also be expressed as a percentage or factor of the mean, and it is calculated as follows:

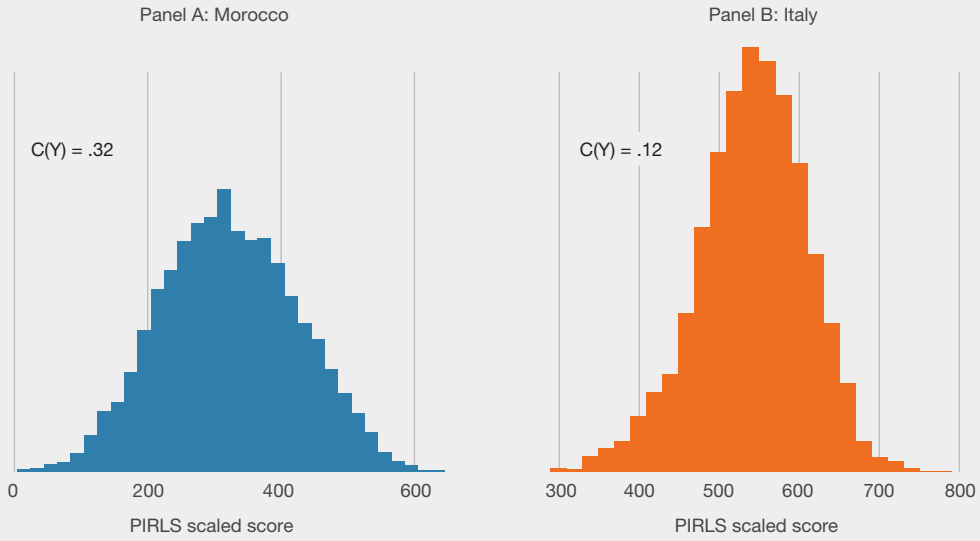
$$C(Y) = SD(Y)/\overline{Y} \qquad [5]$$

The coefficient of variation, unlike the standard deviation, is dimensionless since both the numerator and denominator are measured in the same unit. Further, because of the manner in which the coefficient of variation is constructed, applying a location transformation to the data does not change the perceived level of inequality as measured by this method. **Figure 3.10** plots histograms from PIRLS test score results in Morocco and Italy. From visual inspection, it is clear that Italy's histogram is narrower, indicating a lower degree of inequality in test score performance, whereas Morocco exhibits a larger amount of dispersion in performance. The coefficient of variation in Morocco is 0.32, meaning that one standard deviation is almost 32% of the mean. On the other hand, Italy yields a coefficient of variation of only 0.12, i.e. the standard deviation is only 12% of the mean, indicating a relatively narrower cluster of test scores.

Despite its advantages, the coefficient of variation reintroduces the problem of using ratios in general, namely that the magnitude of the denominator can adversely affect the relative magnitude of the standard deviation. As with any ratio, when the denominator decreases the ratio rises proportionally. In cases where the mean is close to zero, the resulting coefficient of variation will be quite large in magnitude, regardless of the magnitude of the standard deviation. To illustrate the problem, we analyse data on years of schooling in Ethiopia and Burundi, where the former has a higher mean but both have similarly sized standard deviations.[16] Mean years of schooling in Ethiopia is reported to be around 6.7 years with a standard deviation of 4.1, while Burundi reports around 2.8 mean years of schooling and a standard deviation of 3.9 years. The coefficient of variation is 0.6 for Ethiopia and 1.4 for Burundi. With two similarly shaped distributions, a researcher could still encounter vastly different coefficients of variation, and thus different assessments of inequality.

16 The data for Ethiopia and Burundi are from the Demographic and Health Surveys (DHS) administered in 2011 and 2012, respectively.

Figure 3.10 PIRLS test scores from Morocco and Italy



Source: Progress in International Reading Literacy Study (PIRLS), 2011.

3.3.1.4 Cumulative information

Gini coefficient

The Gini coefficient is one of the more popular measures of equality of condition and develops naturally from analysing cumulative distributions. In essence, it aggregates all the gaps between individuals' educational outcomes in a given population and produces a single number, or ratio, to represent the aggregate-level educational inequality. The Gini coefficient is a measure that ranges between 0 and 1, where 0 indicates perfect equality, i.e. all individuals have the same educational condition; whereas 1 indicates perfect inequality, meaning that the entire stock of an educational outcome belongs to one individual. It is calculated by dividing the area between the Lorenz curve and the 45-degree line by the area under the 45-degree line. **Figure 3.11** plots the Lorenz curve for a hypothetical distribution of years of schooling and annotates the definition of the Gini coefficient accordingly.

The Gini coefficient is calculated as follows:

$$G(Y) = \frac{A}{A+B} \quad [6]$$

where $G(Y)$ is the Gini coefficient derived from the distribution of the educational outcome, Y. The area under the 45-degree line (*A*+*B*) and the area between the Lorenz curve and the 45-degree line (*A*) are calculated as follows:

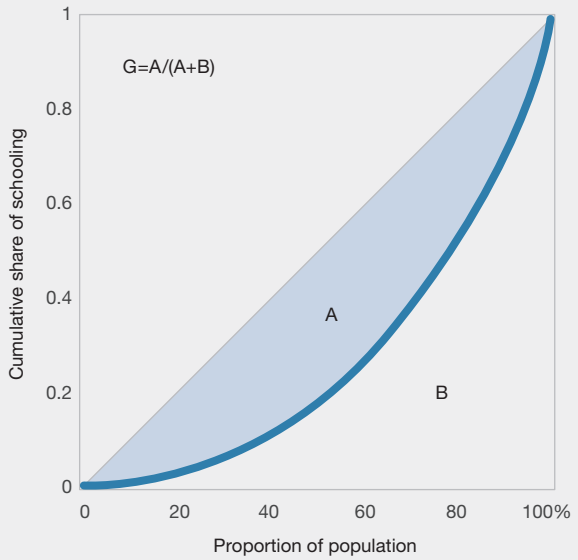
The area under the 45-degree line is the area of a right-angled triangle, as such, the area of *A*+*B* is 0.5 multiplied by the height of the triangle and by the base of the triangle – in this case the height and base of the triangle is equal to 1.

$$A + B = \frac{1}{2} \quad [6a]$$

The area *A* can be calculated as the difference between the area *A*+*B* and the area under the Lorenz curve (*B*), as follows:

$$A = (A + B) - B$$
$$= \frac{1}{2} - \int_0^1 L(Y)\, dY \quad [6b]$$

Figure 3.11 Hypothetical Lorenz Curve for years of schooling and Gini coefficient



where the area under any curve is calculated as the integral of the function that generates that curve, in this case the Lorenz curve. As a result, the Gini coefficient becomes:

$$G(Y) = 1 - 2\int_0^1 L(Y)\,dY \qquad [7a]$$

The computation of integrals is normally reserved for cases when the variable of interest is continuous and exhibits infinitesimally small increments in value. Empirically, most, if not all, educational outcomes are discrete in nature, allowing the Gini coefficient to be calculated as the mean difference between every pair of individuals in a population, divided by the mean (μ) of the variable.

$$G(Y) = \frac{\Sigma_i^n \Sigma_j^n |Y_i - Y_j|}{2n^2\mu} \qquad [7b]$$

Computing a Gini coefficient manually can be cumbersome and use of a specific software is recommended. Alternatively, when ordering the data in ascending value, the Gini coefficient can be calculated in a simpler fashion, which is feasible with any basic spreadsheet application:

$$G(Y) = \frac{\Sigma_i^n (2i - n - 1)\, Y_i}{n^2\mu} \qquad [7c]$$

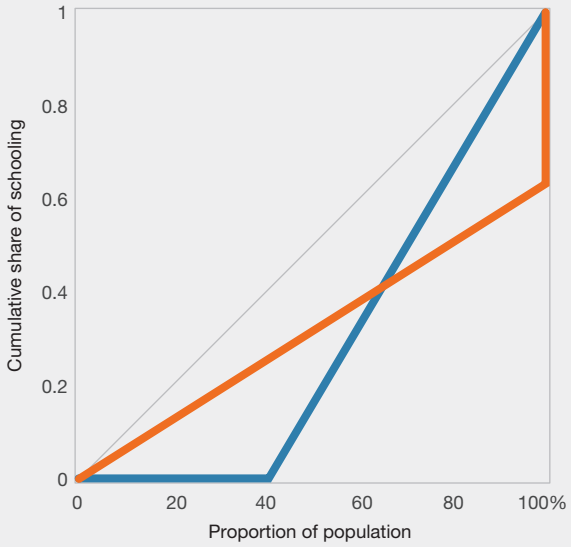
In this case, rather than computing the difference between each pair of individuals of a population resulting in $n(n-1)$ differences that need to be computed, sorting the data in ascending order would only require knowledge of each individual's value of Y, their rank in the population as dictated by their value of Y, the total sample size, and the mean of Y.

It is clear that the Gini coefficient offers several advantages, especially for summarising the state of equality of condition within a population. Like measures of dispersion (standard deviation, variance, coefficient of variation, mean absolute deviation), the Gini coefficient relies on the Lorenz curve, which itself uses all data points from a given distribution and distils it into a single summary statistic that can be easily interpreted. The Gini coefficient also makes it possible to compare levels of inequality across populations on the same scale, regardless of the values of educational outcomes and the population size of each. Last, the Gini coefficient can be decomposed to inform research about the extent of between-group and within-group inequality, which can shed light on the source of the overall inequality and pave the way for a more systematic analysis of equality of condition.[17]

Despite these advantages, the Gini coefficient suffers from certain drawbacks. Because it is an aggregate measure of inequality, it can be insensitive to changes at the tails of the distribution, where the focus of most inequality analyses lies. Suppose we have two independent populations with symmetric educational attainment distributions that produce two Lorenz curves, as shown in **Figure 3.12**. Their curves

17 Group level measures are discussed in detail in Section 3.2.2.

Figure 3.12 Symmetric Lorenz curves for two independent populations with equal Gini coefficients



Note: Hypothetical data.

intersect with the first one (blue line) exhibiting a higher degree of inequality in the bottom half of the distribution while the second one (orange line) exhibits a lower degree of inequality in the bottom half of the distribution and vice versa. Based on the calculated variances, the first is more equitable in terms of educational attainment as the variance is lower. However, in this case, the Gini coefficient for both populations is in fact the same, measured at 42%.

Finally, for an empirical application of the Gini coefficient, we revisit Figure 3.6, which measured the degree of inequality in years of schooling completed in the United States and Burundi. We see that the United States exhibits a low degree of inequality, as the Lorenz curve is within close proximity to the perfect equality line, while that of Burundi is closer to "perfect inequality". The corresponding Gini coefficients for the United States and Burundi are 9.7% and 64%, respectively.

Theil index (generalised entropy)

The final measure of equality of condition discussed in this chapter is the Theil index, which is a special case of the generalised entropy (GE) index. The GE index is a measure of equality of condition derived from information theory that measures the amount of redundancy in the available data. This is an intuitive measure of inequality/equality since perfect equality refers to a scenario where all data points exhibit the same outcome value, which means that the data exhibit a maximum amount of redundancy. In addition to measuring redundancy, the GE index also measures diversity found among the population on a certain outcome. Mathematically, the GE index is calculated as follows:

$$GE\,(\alpha) = \frac{1}{n\alpha\,(\alpha - 1)} \Sigma_i^n \left[\left(\frac{Y_i}{\overline{Y}}\right)^{\alpha} - 1\right] \qquad [8a]$$

where α is the weight given to distances between educational outcomes at different portions of the distribution. Higher values of α will yield an inequality measure that focuses on the upper tail of the distribution and lower values focus on the lower tail, and n is the total sample/population size. The Theil index is a special case of the GE index where α equals 1 and, as such, it is calculated as follows by replacing the value of α with 1:

$$GE\,(1) = T(Y) = \frac{1}{n} \Sigma_i^n \frac{Y_i}{\overline{Y}} \ln\left(\frac{Y_i}{\overline{Y}}\right) \qquad [8b]$$

We can see from equation [8b] that as the value of Y_i increases, or deviates from the mean, so too does $T(Y)$ indicating an increase in inequality. This also means that the value of $T(Y)$ is essentially unbounded. As Y tends to infinity, converges to a maximum value of $\ln\,(n)$ and thus increases with the population size. The Theil index can therefore be difficult to interpret, as it is dependent on the scale of the educational outcome and does not provide any reference to the absolute degree of inequality. This difficulty can be circumvented by dividing equation [8b] by the natural logarithm of the population size, which bounds the Theil index to take on values strictly between 0 and 1. This adjustment can be referred to

as the standardised Theil index, thus rendering the inequality measure comparable across populations regardless of population size, distribution, or shifts therein:

$$\widehat{GE}(1) = \hat{T}(Y) = \frac{1}{\ln(n)} * \frac{1}{n} \Sigma_i^n \frac{Y_i}{\overline{Y}} \ln\left(\frac{Y_i}{\overline{Y}}\right) \quad [8c]$$

The Theil index, or any form of the generalised entropy index, provides several advantages that make it attractive. The index is computationally simple because its calculation requires only individual data points, the mean, and the population size. This makes the Theil index flexible and easy to use because it can be calculated using basic spreadsheet applications such as Excel. Similar to the Gini coefficient and variance-based measures, the Theil index makes use of all observations in a dataset to inform the level of inequality. Moreover, standardising the index enables comparisons across populations. The Theil index can also be decomposed into two components – within-group inequality and between-group inequality.

However, unlike the Gini coefficient and even after standardisation, the value of the Theil index does not have an intuitive interpretation beyond the notion that values closer to 1 indicate high inequality and vice versa. To illustrate the Theil index empirically, we continue with the example of degrees of inequality in Burundi and the United States from the previous section. Recall that the Gini coefficients in Burundi and the United States are 9.7% and 64%, respectively. The Theil index is 0.002 for the United States and 0.092 for Burundi, which shows clearly that Burundi is more unequal than the United States. However, it is difficult to ascertain whether the absolute level of inequality in Burundi is high, since 0.092 is closer to 0 than to 1.

3.3.2 Impartiality

Following up on the discussion of equality of condition using univariate analyses of individual-level outcomes and data, in this section we discuss analyses of impartiality when focusing on bivariate associations with educational outcomes via aggregated data. It is common practice in applied policy research to analyse populations with reference to their social, cultural and economic characteristics. This type of aggregation enables researchers and policymakers to effectively target systematic disparities that are more pronounced among specific groups. A common example of this would be implementing social welfare policies that target the poor, in which case it would be useful to stratify the analysis and examine potential impacts on different socioeconomic groups. Aggregating data allows researchers to examine disparities at the group level using outcomes that are binary at the individual level, such as progression between grades, completion of a level of education, degree attainment, and literacy, among others. This is possible because the mean of a binary variable (coded as 0 and 1) is the proportion that attain a given outcome. As such, analyses of impartiality can be more comprehensive because educational outcomes can be disaggregated or grouped along several demographic, socioeconomic, cultural and other characteristics of the population.

This section discusses equality of condition using aggregated measures that can show whether the distribution of an educational outcome is impartial across different groups. From an empirical standpoint, aggregated data can yield different results for measures of equality because they summarise individual information into a single number for each group. For instance, a common method of analysing the male-female educational attainment gap is to first compute the mean attainment levels by group and then to compare them. However, other information, such as the distribution of attainment, is omitted.

3.3.2.1 Differences

In a group setting, measuring impartiality can be simplified if groups can be defined according to observed characteristics, such as gender, socioeconomic status, ethnic or racial group, migrant status, language spoken at home, and disability. Researchers can then compare the mean (or any

other measure of central tendency) educational outcome between the different groups. In the context of computing between-group differences, the calculation remains unchanged from equation [1], with the exception of the subscripts indexing groups rather than individuals:

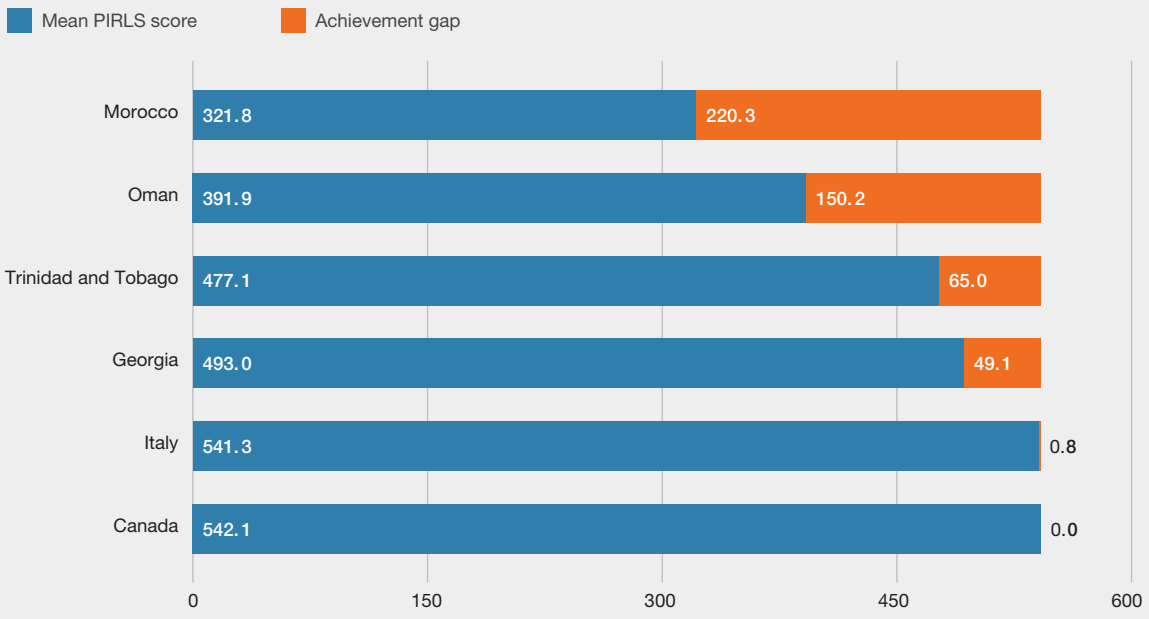
$$D^* = \overline{Y}_i - \overline{Y}_j : i,j = 1,2,3, \ldots, J \quad [9]$$

where D^* denotes the group difference between group i and group j.[18] This is a popular measure, especially when analysing gaps in outcomes between different groups, defined by gender, ethnicity, socioeconomic status or other characteristics.

The advantages and disadvantages of using this measure for group-level aggregate data are similar to those when using individual-level data. The difference or gap is still sensitive to outliers in a group setting but to a lesser extent. Differences and gaps are scale dependent, as the manner in which the educational outcome of interest is measured can influence the magnitude of the gap. In addition, a difference or gap is usually measured between two groups at a time, but it is also feasible to compare several groups by using a reference group. However if researchers were constrained to provide a single statistic to quantify impartiality in a given outcome, there could be a substantial loss of information if there are more than two groups that comprise the whole population.

A common approach to calculating several differences is to measure gaps between every group and a reference group, rather than between every pair of groups. This reference group could be the group with the highest mean achievement, for instance. **Figure 3.13** provides an empirical application of achievement gaps using PIRLS data from Canada,

Figure 3.13 Between-country PIRLS achievement gaps



Source: Progress in International Reading Literacy Study (PIRLS), 2011.

18 From this point forward, group-level statistics will be denoted with an asterisk as a superscript.

Georgia, Italy, Morocco, Oman, and Trinidad and Tobago. Canada exhibits the highest mean score among the countries, while Morocco exhibits the lowest mean score. The achievement gap between Morocco and Canada is 220.3 points. Achievement gaps can be compared between different groups to determine whether the gap is relatively larger or smaller in comparison to other countries. However, gap measures remain problematic because the unit of measurement can be ambiguous, especially when analysing assessment-based scores.

3.3.2.2 Ratios

Measuring impartiality using ratios is especially popular when investigating disparities between two groups and when the scale of the indicator allows it.[19] The formula for calculating the ratio between two groups is the same as equation [2], with subscripts now representing groups rather than individuals:

$$R^* = \frac{\overline{Y}_i}{\overline{Y}_j} \quad [10]$$

R^* denotes the ratio between the mean educational outcomes for group i and group j. Like differences, ratios can only be calculated between two groups at a time. Ratios show disparity between two groups but do not summarise the degree of inequality. Nonetheless, ratios are informative when dealing with binary groups, such as male and female, and when comparing several groups with a reference group.

As an empirical example, we use UIS data to compute the female-to-male out-of-school rate among primary-school-age children across regions. This enables us first to determine the level of impartiality in terms of school access within each region, and second to compare the degree of inequality between regions. The female-to-male ratio of an indicator is commonly called the gender parity index (GPI) and gender parity is typically considered achieved at GPI values between 0.97 and 1.03.

Figure 3.14 plots the adjusted GPI for the primary out-of-school rate in the regions used for monitoring of the SDGs.[20] In six regions – Central Asia, Eastern and South-Eastern Asia, Southern Asia, Northern Africa and Western Asia, sub-Saharan Africa, and Oceania – girls of primary school age are significantly more likely to be out of school than boys. By contrast, boys are more likely to be out of school than girls in two regions, Europe and Northern America, and Latin America and the Caribbean.

3.3.2.3 Dispersion

As explained earlier (Section 3.2.1.3), measures of dispersion are most powerful when examining variability with large sample sizes since the variance is inversely proportional to the size of the sample/population. In addition, measures of variance are most informative under normality conditions, which enables us to apply Chebychev's rule to contextualise the magnitude of a standard deviation. The between-group variance, standard deviation and coefficient of variation are calculated as follows:

$$V^*(Y) = \frac{1}{J}\sum_{j=1}^{J}(\overline{Y}_j - \overline{Y})^2 \quad [11a]$$

$$SD^*(Y) = \sqrt{\frac{1}{J}\sum_{j=1}^{J}(\overline{Y}_j - \overline{Y})^2} \quad [11b]$$

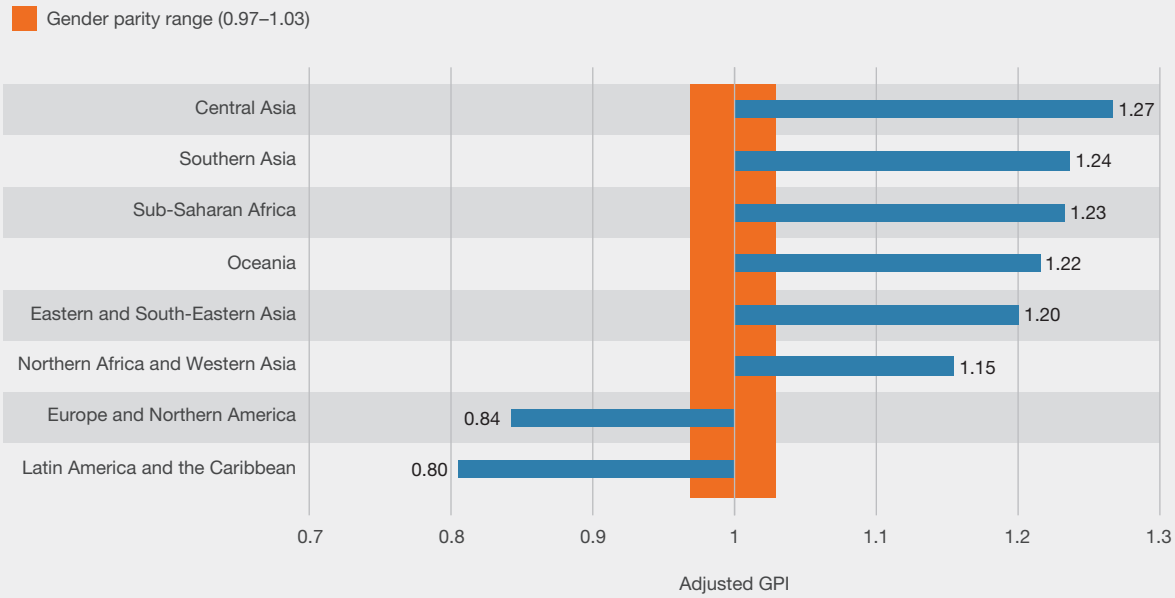
$$C^*(Y) = SD^*(Y) / \overline{Y} \quad [11c]$$

where V^* is the between-group variance, SD^* is the between-group standard deviation, C^* is the between-group coefficient of variation, $\overline{Y}_j$ denotes the mean for group j, and J denotes the total number of groups.

19 For instance, it would not be appropriate to calculate ratios from non-linear scaled scores.

20 The unadjusted GPI is calculated by dividing the female value of an indicator by the male value. The adjusted GPI uses the following methodology: when the ratio of female to male values is less than or equal to 1, the adjusted GPI is identical to the unadjusted GPI. By contrast, when the ratio is greater than 1, the adjusted GPI is calculated by subtracting the male-to-female ratio from 2 (UIS, 2010).

Figure 3.14 Adjusted gender parity index of the primary out-of-school rate, 2016



Source: UNESCO Institute for Statistics database.

Aggregating individual-level data to group-level means has two consequences on the computation of a between-group variance. Computing group-level statistics will, by construction, lead to a significantly diminished sample size because it is safe to assume that J (the number of groups) is substantially smaller than n (the number of individuals). The first consequence is that the tails of the disaggregated distribution are trimmed, thus resulting in the group-level distribution of the educational outcome becoming narrower, which would lead to a smaller level of dispersion in the numerator of the variance formula. The second consequence of using aggregate means is that it introduces small-sample bias since all measures of variance are inversely proportional to n, or in this case, J, which would lead to a larger variance because of a smaller number in the denominator of the variance formula (equation 11a). This results in the group-level variance measure producing two conflicting effects of aggregation where it is unclear which is stronger. Nonetheless, the use of variance as a measure of dispersion introduces an interesting dynamic that decomposes the individual-level variance into a between-group and within-group component as follows:

$$V = \frac{(J-1)\ V^b + (N-J)\ V^w}{N-1} \quad [12]$$

where V^b is the between-group variance, V^w is the within-group variance, and J and N represent the number of groups and the total sample size, respectively. Based on this decomposition, we can measure between-group, within-group and overall variances in one step. This enables a researcher to determine both equality of condition and impartiality through variance by grouping data according to known demographic divisions. Applying this decomposition to the standard deviation and the coefficient of variation yields the same result.

As an example, we return to the data from Figure 3.10 in the previous section that examines the coefficient

of variation for PIRLS test scores in Morocco. We generate an arbitrary grouping of the students who sat for the PIRLS assessment, divide them into five groups, calculate the mean PIRLS test score for each group, and then compute the between-group variance. The overall mean PIRLS score for Morocco is 321.8 points with a standard deviation of 102.1. Applying the variance decomposition method in equation [12], we find that the between-group standard deviation is only 1.29 points and the within-group standard deviation is 102.09 points. This finding clearly shows that almost none of the inequality of condition in PIRLS outcomes are explained by this arbitrary grouping. Conversely, an extreme example would be to divide the population into PIRLS score quintiles or deciles, and the total variance becomes almost completely explained by ability groups.

3.3.2.4 Cumulative information

The Gini coefficient is computed from group-level data in a similar manner to the individual-level index in equation [7b], replacing the sample size with the number of groups and weighting the group data by the group population shares as follows:

$$G^{*}(Y) = \frac{\Sigma_{i}^{n_i} \Sigma_{j}^{n_j} n_i n_j \left| \overline{Y} - \overline{Y}_j \right|}{2 n^2 \mu} \quad [13]$$

G^{*} is the between-group Gini coefficient, i and j represent distinct groups, n_i and n_j represent the number of individuals in each group, n is the overall sample size, and μ is the overall mean of Y. Intuitively, the Gini coefficient group measure is interpreted in the same manner as the individual-level Gini coefficient, with values close to 0 representing a high level of equality and values closer to 1 representing the converse. The Gini coefficient is useful as a relative measure of impartiality, rather than an absolute one. Since the scale is consistent across populations, one can assess which groups exhibit relatively higher degrees of impartiality within the population. However, on an absolute scale, the group-level Gini coefficient suffers from small-sample bias, as in the case of group variances. Again, the group Gini metric relies on group means for the data points used in calculating the Gini coefficient, which results in a narrowed distribution of educational outcomes. The corollary here is that the group-level Gini coefficient is necessarily smaller than the overall Gini coefficient.

The overall Gini coefficient can be shown to be a combination of the between-group and within-group Gini coefficients as follows:

$$G(Y) = G^{*}(Y) + \Sigma s^{k} G^{k} + R \quad [14]$$

where G and G^{*} are the overall and between-group Gini coefficients, respectively. G^{k} represents the within-group Gini coefficient for group k, and s^{k} is the share of total education outcomes owned by group k. Last R, is the residual term that denotes the difference between the sum of the between- and within-group Gini coefficients, and the overall Gini coefficient. In this case, the decomposition of the Gini coefficient enables us to discern the degree to which (in)equality of condition can be explained by impartiality along the groups used for analysis.

As an example, we analyse inequality in completed years of schooling in Malawi. Using data from a DHS administered in 2015, we group individuals in the sample by wealth quintile. The results show that the between-wealth group Gini coefficient is 16.9 points, the within-group Gini coefficient 61.4 points, and the residual is -15.9 points. The overall Gini coefficient of educational attainment in Malawi is 62.4 points, which means that the between-wealth group disparities in years of schooling represent about 27% (16.9 divided by 62.4) of the total disparity. However, it is important to note that the presence of the residual term can make the interpretation somewhat challenging since the within-group disparity, in this case, is approximately 98% of the total disparity (61.4 divided by 62.4). Clearly, the proportions of total inequality explained by the within-group disparity and the between-group disparity sum to more than 100%.

The Theil index, on the other hand, circumvents this disadvantage because the overall Theil index decomposes additively between the between-group and within-group inequalities without the presence of a residual term to maintain the whole. The mathematical representation of the Theil index decomposition is as follows:

$$T(Y) = \sum s_k \frac{\overline{Y}_k}{\overline{Y}} T_k + T^*(Y) \qquad [15]$$

where

$$T^*(Y) = \sum_k s_k \frac{\overline{Y}_k}{\overline{Y}} \ln\left(\frac{\overline{Y}_k}{\overline{Y}}\right) \qquad [16]$$

T_k is the within-group Theil index for group k, T^* is the between-group Theil index, and T is the overall Theil index. In this case, it is much easier to see how the within- (equality of condition) and between-group level (impartiality) inequalities combine and enable researchers to ascertain the degree to which specific subgroup divisions can explain the overall level of equality of condition within a population.

Using the same data from Malawi grouped by wealth quintile, we compute the between- and within-group Theil indices for the years of schooling completed by individuals. We find that the between-group Theil index is 0.048, the within-group Theil index is 0.787 and the overall Theil index is 0.835. The result suggests that wealth-based divisions of the population explain 5.7% (0.048 divided by 0.835) of the overall inequality.

3.3.2.5 Analytic tools for testing impartiality

To measure and quantify impartiality of educational outcomes among subgroups, we introduce three common statistical tools: the simple Pearson correlation coefficient, F-tests (equality across several groups), and the coefficient of determination (R^2 or R-squared).

The correlation coefficient is a measure of impartiality when analysing bivariate associations between one continuous educational outcome and one continuous quantification of population characteristics. For instance, if we want to assess whether educational attainment (measured in years of schooling) is associated with household wealth, we can correlate the two measures and the resulting coefficient indicates whether the educational outcome is independent of wealth. Moreover, a correlation coefficient of zero on such a bivariate association would indicate that the two constructs are uncorrelated, i.e. impartial.

In cases where the grouping factor or population trait is not continuous, binary or categorical, we can rely on a simple ANOVA or F-test to determine independence between the two. Formally, the F-test is used to determine whether the mean educational outcomes across all groups are equal (impartial). For example, if we want to determine whether the rate of out-of-school children ($\overline{Y}$) is the same across wealth quintiles, the hypothesis test is set up as follows:

$$H_0 : \overline{Y}_{Q1} = \overline{Y}_{Q2} = \overline{Y}_{Q3} = \overline{Y}_{Q4} = \overline{Y}_{Q5} \qquad [17]$$

where Qi refers to the i^{th} quintile in a given sample or population, and the null hypothesis tests whether the mean outcome is equal across all quintiles. The alternative hypothesis would be that at least one of the quintiles has a mean out-of-school rate that is not equal to the others. The advantage of this test is that it is simple to use and can be computed using any standard statistical software. In addition, the data requirements are not as restrictive as for the Pearson correlation coefficient, as the F-test can be used with categorical and continuous data for both the educational outcome and the bivariate association.[21]

Last, we discuss the R-squared, which is a multivariate version of the Pearson correlation

21 Note that F-tests can be augmented to the Chow test, which is a partial F-test from a linear regression, where group membership can be tested in the presence of control variables, thus making the test useful in a multivariate environment.

coefficient used in linear regression analyses. The R-squared enables the same mode of analysis as the simple correlation coefficient to test for impartiality in a multivariate regression analysis. An R-squared of zero (or close to zero) would suggest that the distribution of the educational outcome of interest is not associated with the characteristics of individuals included in the regression analysis.

However, the R-squared coefficient has certain drawbacks. Although it can be computed for any type of data, its use with binary or categorical outcomes suffers from the same problems as group-based variances. The variances would be deflated, since possible values are only zero or one and would lead to smaller values of the R-squared coefficient when there might be a strong statistical association between the dependent (outcome) and explanatory variables. In addition, the R-squared can be a flawed test since it increases with the number of right-hand-side variables included in the regression analysis. In both cases, the analysis is susceptible to both Type I and II errors, i.e. incorrect rejection of a true null hypothesis or incorrect retention of a false null hypothesis.

3.4 AVAILABILITY AND COMPARABILITY OF EDUCATION DATA ON EQUITY DIMENSIONS

That education in a society should be equally distributed and impartial, regardless of individual or group characteristics, is a central tenet of educational equity. The group measures introduced earlier provide ways to assess the impartiality of education systems. But they require thoughtful identification of the characteristics that are likely to make children and young people more vulnerable and, therefore, more likely to undermine impartiality and ultimately equity in education.

Although the determinants of disadvantage vary by context, certain factors have emerged in international frameworks that seek to improve equity in education. For example, the Incheon Declaration (UNESCO, 2016a) identifies the following key dimensions that need to be taken into account to achieve equity: gender, disability, forced displacement, and diversity along cultural, linguistic and ethnic lines. In addition, poverty, residency, gender and disability are named in SDG 10 on reduced inequality and in Goal 4 on education. These characteristics are all often associated with resource deprivation or discrimination and have known predictive effects on education experiences and outcomes. Section 4.2 summarizes some of the disparities linked to personal and household characteristics.

Although international frameworks make clear that the education agenda must target *all marginalised individuals and groups*, more efforts are needed to collect disaggregated data to address different equity dimensions. It is important to note that characteristics not explicitly mentioned in the SDGs also play a role in equity. For example, individuals' gender identity can strongly influence their educational experiences but such information remains largely absent from many data collection efforts.

The following sections explore how and whether equity dimensions are currently being measured in major international and regional sources of education statistics. The results show that, at present, the potential to assess the impartiality of education systems is constrained by the scarcity and non-comparability of data on key equity dimensions. Although some promising efforts are underway, greater investment is needed to ensure better and more widespread equity analyses in the future.[22]

22 For more detailed reviews on the availability and comparability of data for the measurement of education equity, see Education Equity Research Initiative (2016) and UIS (2016).

3.4.1 Scarcity of data on key dimensions of equity

Table 3.3 documents the availability of education data by equity dimension across popular international and regional education databases. As Table 3.3 shows, education data disaggregated by sex are more widely available than by other dimensions. This is due to the relative ease of collecting data by sex through different modes of data collection, such as school censuses, household surveys, population censuses or learning assessments. Indeed, measurement of education equity with respect to gender differences tends to be limited more by the unavailability of certain education data (e.g. on topics such as learning or absenteeism) than by the inability to disaggregate available statistics by sex, as recognised by the MDGs and now by the SDGs.

Nevertheless, gaps in the availability of gender-disaggregated education statistics persist. A recent UIS survey found that only 85% of existing data for SDG 4 indicators on education can be disaggregated by gender (UIS, 2016). More and better data are also needed on the intersection between gender and other dimensions of inequality. Such intersection often compounds educational vulnerability, with poorer or more rural girls at particular disadvantage in some contexts (see Section 4.2.1). These intersecting dimensions can be measured from certain primary sources (e.g. large-scale household surveys) and are available from some databases, including the UIS. Stat online database, the World Inequality Database on Education (WIDE) and the Education Policy and Data Center (EPDC), as well as in the 2016 *Global Education Monitoring Report* (UNESCO, 2016b). However, improving their coverage – and coverage of gender-disaggregated education statistics in general – must remain a focus in present and future work.

Residency (urban or rural, sub-national region) is relatively simple to document, like gender. In fact, all primary data sources covered in Table 3.3 include urban-rural residency as an equity dimension.[23] By comparison, the coverage of sub-national regions is more common in household-based surveys and school surveys than in school-based student assessments, which often have smaller samples and are limited in the extent of detail they can report. While residency is widely collected as an equity dimension in primary sources, it is not always shared as an equity dimension in major international education databases. Urban-rural residency is included in 7 of 11 international databases, but sub-national regions in just 3 databases *(see Table 3.3).* One constraint on greater international coverage of residency information stems from the lack of comparability, an issue that will be further explored in the following section. In short, while residency is regularly collected by primary sources, its utilisation at the international level is sporadic at best.

Whereas collecting gender and residency data is straightforward, it is more difficult in school-based data collection (e.g. school censuses) to gather information about other important equity dimensions – such as wealth, ethnicity, language, disability and migration status.[24] Indeed, as Table 3.3 shows, apart from school censuses, all primary sources collect information on wealth or socioeconomic status, though only one-half of international and regional databases publish education statistics using wealth as an equity dimension. Information on ethnicity and language is also collected by most primary sources, at least in some countries, but

23 Whether residency data are collected by national administrative sources varies by country and over time. Such data are often but not always collected by school censuses and other national sources.

24 Although collecting information about these individual- and household-level characteristics at schools poses difficulties, it is feasible to do so. For example, UNICEF (2014) provides recommendations for measuring disability in school censuses, and the UIS and the Global Partnership for Education began collaborating in 2017 on the production of statistics on education and disability based on administrative records.

Table 3.3 Availability of equity dimension data in international and regional databases and primary sources

		Equity dimension							Education focus							Education level				
Institution	Source	Sex	Residency - urban/rural	Residency - region	Wealth or SES	Ethnicity, religion, or language	Disability	Migration status	Resources/inputs	Access	Participation	Retention/survival	Attainment	Learning outcomes	Literacy/skills	Pre-primary	Primary	Lower secondary	Upper secondary	Post-secondary
International databases																				
EPDC - FHI 360	epdc.org	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ILO, UNICEF, World Bank	Understanding Children's Work	✓	✓								✓									
OECD	OECD.stat	✓							✓	✓	✓	✓	✓			✓	✓	✓	✓	✓
UIS	UIS.Stat - administrative data	✓							✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
UIS	UIS.Stat - household survey data	✓	✓		✓						✓						✓	✓		
UN	UNGEI	✓									✓	✓			✓		✓	✓		
UNESCO GEMR	WIDE	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
UNICEF	data.unicef.org	✓	✓		✓					✓	✓				✓		✓			
UNICEF	TransMonEE *	✓					✓				✓					✓	✓	✓	✓	
USAID	Early Grade Reading Barometer	✓	✓	✓										✓			✓			
World Bank	EdStats/Educational Equality	✓	✓		✓				✓	✓	✓	✓	✓	✓			✓	✓	✓	
Household-based surveys and assessments																				
ICF International	DHS	✓	✓	✓	✓	✓				✓	✓	✓	✓		✓		✓	✓	✓	✓
National agencies	Population censuses	varies by country							varies by country							varies by country				
PAL Network members	PAL Network learning assessments	✓	✓	✓	✓	✓				✓	✓			✓			✓			
UNICEF	MICS	✓	✓	✓	✓	✓				✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
School-based student assessments																				
CONFEMEN	PASEC *	✓	✓		✓	✓	✓		✓					✓			✓			
IEA	PIRLS and TIMMS	✓	✓		✓	✓			✓					✓			✓	✓		
OECD	PISA	✓	✓	✓	✓	✓			✓					✓					✓	
SACMEQ	SACMEQ *	✓	✓		✓	✓			✓					✓			✓			
UNESCO/LLECE	PERCE, SERCE, TERCE *	✓	✓		✓	✓			✓					✓			✓			
School- or teacher-level surveys																				
National agencies	Annual school census	varies by country							varies by country							varies by country				

* Regional initiatives

Source: Adapted from UIS (2016) and Education Equity Research Initiative (2016).

is only published regularly in the WIDE database. As with residency, challenges with comparability of data on these topics, especially ethnicity and language, limit more widespread use. For disability and forced displacement, primary data collection is very limited even though these dimensions are widely recognised as factors that are likely to create educational disadvantage. It is encouraging that international efforts led by the Washington Group on Disability Statistics (subsequently referred to as the Washington Group) and UNICEF have championed measurement of disability, with more widespread collection of disability data likely to be realised in the near future. The 2016 *Global Education Monitoring Report* highlighted the need for – and challenges of collecting – data on forced displacement (UNESCO, 2016b), which will hopefully garner more attention for measurement on that dimension.

In sum, national agencies and international organizations recognise the importance of being able to disaggregate data by key equity dimensions. In practice, however, data are more widely available for some dimensions (e.g. gender, residency and wealth) than others (e.g. ethnicity, language, disability or migration status). In some cases, like residency, data availability is limited more by problems of comparability and availability at the international level than by data collection itself. For other dimensions, notably disability and migrant status, there are limited data available from primary sources. An important next step is therefore to mainstream the collection of data on those topics. The next section turns from the topic of data availability to the specific challenges that arise with data comparability when using key equity dimensions.

3.4.2 Comparability of data on key dimensions of equity

Even where data on equity dimensions are available, the differences between definitions, modes of data collection and culturally-specific response options make it difficult to make comparisons within and across countries. This section explores these challenges further and provides examples from the primary data sources shown in Table 3.3.

One type of comparability issue arises from differences in how dimensions are defined across sources. For example, there are often country-specific definitions of urban and rural that are used by DHS and other surveys, with some based on population size and others based on infrastructure (DHS, 2013). Definitions of wealth also vary, and the majority of sources include a composite index of relative economic or socioeconomic wellbeing, though the specific items included in these measures vary by data source, country and time. While such discrepancies need not prevent the use of these dimensions in international comparisons, they do mean that comparisons should be understood to represent country-specific, and sometimes data-source-specific, definitions of concepts.

Comparability can also be undermined by differences in categories of responses, especially with measures of identity, such as ethnicity, religion, indigenous group or language. Identity groups are important for equity analyses, because disparities between them may suggest discriminatory policies or unfair resource allocation. However, comparability of data on identity groups within countries depends on which identity groups are considered. For some data sources, the choice is driven by the relevance of certain social divisions – all countries have some degree of ethnic, religious and linguistic diversity but those cleavages do not always drive disadvantage. In Latin America, for example, indigenous status is an important social division with ramifications for educational inequality, and the Latin American Laboratory for Assessment of the Quality of Education (LLECE) – which has conducted learning assessments across the region, as well as censuses and household surveys – seeks information about it. For learning assessments, language is of particular interest because command of the language of instruction is a strong determinant of academic performance (Brock-Utne, 2007).

Even when ethnicity, religion or language are measured repeatedly over time, the groups identified might vary. For example, 1989 and 1999 censuses in Viet Nam list roughly 50 groups, most of which are common to both surveys, whereas the 2009 census includes only three categories: Kinh (the ethnic majority group), other ethnic group and unknown. Shifting categories result from survey design or changes in the social relevance of certain categories. Modern conceptualisations of ethnicity tend to see ethnic identities as simultaneously malleable and durable, with the social or political relevance of certain categories evolving over time (Brown and Langer, 2010). In quantifying the concept of ethnicity, surveys must wrestle with the challenge of assigning concrete codes to blurry concepts and must negotiate occasional tensions among feasibility of measuring many groups, relevance and comparability over time.

Differences in cultural understandings of concepts may also complicate comparability. How to avoid such problems has been a particular focus in the measurement of disability, as the concept of disability itself is often deeply rooted in cultural norms. Population censuses, for example, may reflect local understandings of disability and priorities for data collection and therefore may not be compatible with international standards. The Integrated Public Use Microdata System (IPUMS), which works to harmonise data from over 270 censuses, cautions that, even when responses on disability can be presented under a common variable, comparability across surveys is complicated by differences in questionnaire phrasing, what counts as a disability (e.g. some censuses include chronic diseases under disability) and how severe a condition must be to be labelled a disability. Fortunately, question sets on disability developed by the Washington Group and UNICEF are increasingly adopted in household survey programmes and have also been endorsed in the census guidelines by the United Nations Statistics Division (UNSD) (UN, 2015) and the Conference of European Statisticians (UNECE, 2015), meaning that availability and comparability of disability data should continue to improve (see Section 4.2.1 for a more detailed discussion of the links between education and disability).

In conclusion, more comparable data on key dimensions of inequality and greater use of such data in equity assessments are achievable. The growing consensus around the measurement of adult and child disability indicates that greater harmony in measurement of key equity dimensions is theoretically possible. International attention to the measurement of other dimensions could yield improved availability and comparability of data for those areas as well. The effects of forced migration on equity in education, for example, warrant special attention. In advocacy literature this dimension is regularly tied to severe vulnerability, yet it has been difficult to measure in many national and international surveys on education to date.

3.5 DESIGN AND SAMPLING CONSIDERATIONS

Having noted the scarcity of available data on equity, we turn to the need for consistency and comparability of generating data across key dimensions of equity and integrating equity measures into every data collection effort of reasonable scale. However, an important challenge facing organizations interested in equity analysis is to determine what constitutes an adequate sample size – one that would allow for disaggregation of data without substantial loss of reliability. Within the context of most programme-level research and evaluation, the desire to have an informative sample that allows for meaningful analysis must be balanced with a realistic budget and a feasible time frame for collection and analysis.

Many texts discuss general parameters for determining appropriate sample size, depending on the assumptions and starting conditions, such as the desired level of precision for the sample estimate and its confidence level, intra-class correlation (for clustered samples) and the level of statistical power. In evaluation settings, the statistical power

parameter is replaced by the magnitude of expected differences between treatment and comparison conditions *(see Hedges and Rhoads, 2010; Gelman and Hill, 2006)*. The EGRA Toolkit (RTI, 2015, p. 154) provides a breakdown of these considerations for studies involving the administration of EGRA assessment data.

Each of these parameters carries implications for the size and, consequently, the cost of a given data collection effort. Conducting quantitative analyses with a focus on equity increases the complexity of the sampling framework needed to achieve adequate statistical power. In this case, sampling and power considerations need to be made for each sub-group of interest. Indeed, desired levels of precision for a given estimate largely drive the sample size and cost: reducing standard error for a given population-based estimate by *one-half* may require a *quadrupling* of the sample size (UN, 2008, p. 35) – an increase most development organizations are not able to handle.

To help researchers and practitioners with an interest in equity-oriented analysis to make decisions regarding sampling for impact analysis, we recommend the following steps:

1. Determine which dimensions will act as domains in your analysis.
2. Determine whether certain intersections of equity dimensions should be treated as domains.
3. Obtain an estimated proportion of each domain or dimension in the population.
4. Determine whether a domain can be purposefully oversampled to gather a sufficient number of observations.
5. Determine acceptable cluster size.
6. In evaluation settings, determine the acceptable minimum group-level effect size.

A general description of these steps follows. More detail is available in recommendations issued by the inter-organizational Education Equity Research Initiative (Omoeva et al., 2017).

Determine which dimensions will act as *domains* in your analysis. Household surveys distinguish between *domains* and *tabulation categories* for sub-population analysis. Domains represent critical subpopulations for the study in question, important enough to justify substantial cost increases required to reach sufficient sample sizes within each domain. By contrast, tabulation categories are informative but do not require the same degree of precision of group-level estimates and therefore are not factored into the sample size estimation.

Determine whether certain intersections of equity dimensions should be treated as domains. With multiple dimensions of equity present in a dataset, choices may have to be made as to which intersections of characteristics will require a closer examination as *domains* of analysis. Intersections combine multiple dimensions within a single category, such as the intersection of gender and poverty, for example. Naturally, an intersection of dimensions creates a smaller cell size within a dataset, therefore requiring higher statistical power – and in most cases, a larger overall sample size – to generate reliable estimates.

Obtain an estimated proportion of each domain or dimension in the population. While it is not always possible to know how many observations for a given domain one can expect to obtain in a sample, existing household surveys and censuses often provide a general gauge of how large a given disadvantaged group is within a general population. One can expect, for example, to find roughly 50% females in a simple random sample of households, and administrative data from a school census will indicate what proportion of girls can be expected to be present in a school sample. Similarly, rough proportions may be obtained from prior sources on the presence of ethnic, religious or linguistic minorities, and persons with certain types of disabilities. While the actual proportion in a given sample may vary, prior information provides a useful starting point for subsequent decision making around

sampling and design. When no such information is available, assumptions will have to be made at first and samples will need to be adjusted once initial data become available.

Table 3.4 illustrates the ways in which intersections of dimensions reduce the effective cell size and, as a result, decrease our ability to reliably estimate a parameter. We may start with an overall expectation of 50% girls, of which 30% may be rural and 20% members of a minority ethnic group. However, if our equity analysis is focusing on rural girls, or girls of a particular ethnic minority, we must account for the fact that this sub-group will be only 15% or 10% of our sample. If either of these intersections is a *domain* for our analysis, the sample size calculation should be run with the smallest domain in mind.

Determine whether a domain can be purposefully oversampled to gather a sufficient number of observations. Since the determination of sample size is largely driven by the smallest domain of interest, it is important to consider whether purposeful oversampling is generally feasible to ensure a sufficient number of observations within a cell. The question of whether a group can be identified *a priori* and targeted for additional sampling can be context-driven – as it would be for ethnic minorities, for example. In other situations, such as those involving many types of disability or displacement, this may not be feasible – in which case it will be necessary to decide if obtaining a sufficient number of observations for that group will require a larger sample.

Determine acceptable cluster size. Another decision needed when determining the sample size and the overall cost of a given data collection effort is the acceptable cluster size. In many cases, data for education research and monitoring purposes are clustered at the school level, with the school serving as the primary sampling unit. Deciding what would be a realistic number of observations that can be gathered on students within a school may provide a useful approach to increasing the sample size and, consequently, the number of observations on a given equity dimension without dramatically increasing the cost of the study, which is largely driven by the number of clusters to be visited during data collection. It is also helpful to determine the maximum cluster size that allows for the optimal number of schools visited within a day, and to compare the resulting number with the number of days and sites allowed by the budget.

In evaluation settings: determine the acceptable minimum group-level effect size. A key ingredient in the decisionmaking process for impact evaluation data collection is the magnitude of change in the outcome that can be attributed to the treatment or intervention in question. Many texts on evaluation methodology emphasise the important research task of determining what magnitude of change is worth searching for, because smaller changes in outcomes

Table 3.4 Expected domain size across three equity dimensions of different proportions

	Population proportion	Dimension 1: Girls	Dimension 2: Rural	Dimension 3: Minority
		50%	30%	20%
Dimension 1: Girls	50%			
Dimension 2: Rural	30%	15%		
Dimension 3: Minority	20%	10%	6%	

are generally more difficult to detect and will therefore require larger sample sizes.[25]

3.6 EQUITY ANALYSIS: A PROPOSED SEQUENCE

In setting up the research with a lens focused on equity, we propose a sequence that first quantifies and visualises equality of condition, then moves on to describing impartiality and, finally, tests for impartiality of programme impacts. We narrate the analytic sequence via a series of research questions that can either be combined into a single study or studied separately.

Identify the equity dimensions of interest. When investigating equality of condition or impartiality in any research context, it is important to first define population divisions that group individuals into demographic categories. Key dimensions of equity are gender, ethnicity/race, residence, poverty, disability and immigration/migration status. Once these key dimensions are identified, conducting simple descriptive analyses to summarise the composition of the sample will help to contextualise any subsequent stratified analyses to investigate pre-existing disparities that may exist between individuals and between groups. From a statistical standpoint, it is important to summarise the equity group distributions to show the group's size and its size relative to other groups. As discussed earlier, all metrics of inequality at the group level are a function of the group's size. There are, however, cases where a minority group can be so small that the inequality metrics do not reveal the true extent of the disparities between the groups.

Summarise observable characteristics by equity dimension. The second step that we recommend is to summarise the groups along their observable characteristics. This step aids researchers in constructing statistical profiles for each group that show discrepancies not only in their educational outcomes but also in their baseline characteristics. This analysis achieves two objectives simultaneously. The first is confirmatory in terms of examining differences, or gaps, in the educational outcomes of interest that exist between different groups (this is also applicable across equity dimensions). The second can be thought of as exploratory, as a basic attempt to identify determinants of the outcome disparities, if any exist. This step does not substitute for rigorous analysis (i.e. estimating conditional expectations by constructing appropriate counterfactuals) of determinants of systematic disparities between groups and individuals, but provides researchers with a first step in identifying possible areas of interest that could explain these gaps.

From a practical perspective, simple summary profiles can be constructed by computing simple means of all outcomes and observable characteristics for each group across equity dimensions. At this point we are now ready to compute differences and gaps, and measure variance. The basic inequality metrics can be used in this case to determine both between-group and within-group inequalities that exist in the data. Finally, simple t-tests or F-tests can be conducted to determine whether the observed differences are statistically significant.

Analyse overall outcome distributions. In the previous step, we discussed how summary statistics describe the overall state of the analytic sample, while also providing an overview of the general state of inequality that may exist at the individual and at the group-aggregate level. The logical next step is to examine the full set of information available by focusing on distributional analyses of the educational outcomes of interest. Here, we recommend relying on the visualisation tools detailed at the beginning of this chapter. For instance, a histogram or PDF similar to Figures 3.1 and 3.2 will describe the overall degrees

25 See Bamberger et al. (2006), Gelman and Hill (2007) and Kish (2004), among others.

of disparity that exist between the educational outcomes of interest.

Analyse outcomes by equity dimension. The analysis of the overall educational outcome distributions can be augmented by overlaying the distributions of the outcome for each equity group of interest. This will show both the absolute degree of inequality that exists between the groups by contrasting the spatial location of the distributions, and the within-group degree of inequality where researchers can examine the range of disparity for each group separately. Last, researchers can also compare the within-group inequality across groups to show not only disparities between groups in terms of mean outcomes, but also to inform the degree of inequality that exists within each group.

Estimate main effects of interest, overall and stratified by equity dimension. The final step in most quantitative studies is to implement more rigorous methods to identify plausible effects of policies, interventions, and other exogenous shifts on outcomes. From an equity perspective it is not enough to identify the overall causal effect of a policy. While this is a very informative and important research finding, inspecting effects on the whole population assumes a certain degree of homogeneity in the effect, in that it will have a similar effect across all individuals and groups. Yet countless empirical studies show this often produces an inaccurate assessment. Our recommended approach, outlined below, allows researchers to test statistically whether the effects are heterogeneous or not.

Researchers can formally introduce heterogeneity into a regression analysis framework by interacting the independent variable (e.g. the treatment variable) with indicators for group membership. This will stratify the overall effect to show whether the "treatment" does in fact affect different groups in a systematically different manner. The last step of this piece of the analysis would be to conduct post-estimation testing on the equity-by-treatment interaction effects to determine whether the estimated effects are statistically heterogeneous.

REFERENCES

Bamberger, M., J. Rugh, and L. Mabry (2011). *Real World Evaluation: Working under Budget, Time, Data, and Political Constraints*. Thousand Oaks: Sage.

Brock-Utne, B. (2007). "Learning through a familiar language versus learning through a foreign language: A look into some secondary school classrooms in Tanzania". *International Journal of Educational Development.* Vol. 27, No. 5, pp. 487-498.

Brown, G. K. and A. Langer (2010). "Conceptualizing and measuring ethnicity". *Oxford Development Studies, 38*(4), 411-436.

Demographic and Health Surveys Program (DHS) (2013). Official DHS response in *The DHS Program User Forum*. Retrieved on 18 March 2013 from https://userforum.dhsprogram.com/index.php?t=msg&th=55&start=0&

Education Equity Research Initiative (2016). *Measuring Equity in Education: Review of the Global and Programmatic Data Landscape*. Washington, DC: Education Equity Research Initiative.

Gelman, A. and J. Hill (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Hedges, L. V. and C. Rhoads (2010). *Statistical Power Analysis in Education Research.* NCSER 2010-3006. Washington, DC: National Center for Special Education Research, Institute of Education Sciences, US Department of Education. https://ies.ed.gov/ncser/pubs/20103006/pdf/20103006.pdf

Kish, L. (2004). *Statistical Design for Research*. Hoboken: John Wiley & Sons.

Omoeva, C., W. Moussa, A. J. Dowd, A. Mulcahy-Dunn, K. Alexander, C. Cummiskey and W. Wadhwa (2017). *Practical Recommendations for Improving Comparability and Reliability of Equity Analysis in Education*. Education Equity Research Initiative. Washington, D.C.: FHI 360.

Pen, J. (1971). *Income Distribution: Facts, Theories, Policies*. New York: Praeger.

RTI International (2015). *Early Grade Reading Assessment (EGRA) Toolkit* (2nd edition). Washington, DC: USAID. https://globalreadingnetwork.net/resources/early-grade-reading-assessment-egra-toolkit-second-edition

United Nations (2015). *Principles and Recommendations for Population and Housing Censuses: Revision 3*. New York: United Nations. http://unstats.un.org/unsd/publication/seriesM/Series_M67rev3en.pdf

United Nations Department of Economic and Social Affairs (UNDESA) (2008). *Designing Household Survey Samples: Practical Guidelines*. ST/ESA/STAT/SER.F/98. New York: United Nations. https://unstats.un.org/unsd/demographic/sources/surveys/Series_F98en.pdf

United Nations Economic Commission for Europe (UNECE) (2015). *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing*. New York and Geneva: United Nations. http://www.unece.org/fileadmin/DAM/stats/publications/2015/ECE_CES_41_WEB.pdf

UNESCO (2016a). "Education 2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4." ED-2016/WS/28. Paris: UNESCO. http://unesdoc.unesco.org/Ulis/cgi-bin/ulis.pl?catno=245656

—— (2016b). *Education for People and Planet: Creating Sustainable Futures for All – Global Education Monitoring Report 2016*. Paris: UNESCO. http://unesdoc.unesco.org/images/0024/002457/245752e.pdf

UNESCO Institute for Statistics (UIS) (2010). *Global Education Digest 2010: Comparing Education Statistics across the World*. Montreal: UIS. http://www.uis.unesco.org/Library/Documents/GED_2010_EN.pdf

—— (2016). *Sustainable Development Data Digest: Laying the Foundation to Measure Sustainable Development Goal 4*. Montreal: UIS.

UNICEF (2014). *Education Management Information Systems and Children with Disabilities: Webinar 6 – Companion Technical Booklet*. New York: UNICEF. http://www.inclusive-education.org/sites/default/files/uploads/booklets/IE_Webinar_Booklet_6.pdf

4. Measuring equity for national education planning

BY BEN ALCOTT, PAULINE ROSE, RICARDO SABATES AND RODRIGO TORRES
Research for Equitable Access and Learning (REAL) Centre, University of Cambridge

INTRODUCTION

This chapter presents evidence on the extent to which different indicators included in national education plans take account of equity as discussed in Chapters 2 and 3. It is fairly common to find equity measures in plans for indicators related to access at the primary level. But few national education plans include indicators for learning and, for those that do so, the main dimension of inequality included is sex. Where plans do include measures of equity, these are most often associated with impartiality in that they track sub-groups of the population separately or assess parity between these sub-groups. This chapter highlights positive country examples for tracking progress to achieve equity in access and learning.

It then highlights the importance of including a wider range of dimensions of disadvantage within education plans, discusses what data need to be collected and proposes methods to track progress to identify how inequalities have changed over time. The chapter aims in particular to advise policymakers on what information should be taken into account when deciding on the types of indicators that are suitable for tracking progress on learning.

4.1 MEASURING EQUITY WITHIN NATIONAL EDUCATION PLANS

Our analysis of national education plans identifies the presence of equity dimensions included in indicators for tracking progress towards access and learning at different education levels. It further looks at how these equity dimensions relate to the indicators set in SDG 4, with a focus on the pre-primary, primary and secondary levels as set out in SDG Targets 4.1 ("by 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes") and 4.2 ("by 2030, ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education").

Taking into account the lessons from Chapters 2 and 3, this chapter acknowledges the gap between the types of approaches that are ideally expected, and the reality of what is currently being adopted by most countries, with a focus on those furthest away from SDG Targets 4.1 and 4.2.

National education plans are the main tool that most governments use to plan and implement their policy agenda. In this chapter we assess the types of indicators being used for measuring learning at different education levels and whether equity elements are included. We draw some lessons from countries currently implementing plans to measure equity in learning and in education.

4.1.1 Methodology

Countries were selected from all regions as defined for SDG monitoring.[26] Countries that did not have a plan since 2004 were omitted.[27] Of the 75 countries identified, 34 were from sub-Saharan Africa, 8 from Northern Africa and Western Asia, 9 from Latin America and the Caribbean, 9 from Central and Southern Asia, 5 from Eastern and South-Eastern Asia, 8 from Oceania, and 2 from Europe and Northern America *(see **Annex B**).* As a starting point, we reviewed 52 national education plans available on the Global Partnership for Education (GPE) website. We then broadened geographical coverage with 23 additional plans, ensuring that examples of education plans in English, French, Portuguese and Spanish were included.[28]

We produced a template to record country data on indicators associated with access to education and learning (related to SDG Targets 4.1 and 4.2) and their characteristics, including information on equity. Recognising that national education plans are organized in different ways, with varying amounts of detail, we adopted a common approach for seeking relevant information to complete the template.[29] The protocol involved the following steps:

1. Review the plan looking at the document structure and contents, identifying the information of interest (or at least part of it). Highlight any relevant information.
2. Search for agreed keywords in the text in order to find out where relevant topics and specific information of interest are located.
3. Read in more detail sections which contain information about goals, targets and indicators for the topics of interest.
4. Highlight those paragraphs or tables of interest which contain relevant information, so that it can be analysed and extracted.
5. Review the text again, focusing on the titles and highlighted text to get a sense of how much of the information that needs to be included in the template is available and what information is missing.
6. Search through the text again using keywords to locate missing information.
7. Complete the template by education level or by type of indicator.

4.1.2 Indicators included in national education plans

Access

Indicators for tracking progress on access to education that are most commonly found in national education plans relate to participation and completion at each education level. In general, indicators for participation included in national education plans relate to the ones in the SDG list of thematic indicators for Targets 4.1 and 4.2[30] but are less comprehensive as they usually focus just on enrolment and completion rates rather than the complete list of indicators detailed in those targets *(see **Table 4.1**)*.[31]

26 Each country was assigned to the corresponding region according to the United Nations SDGs Regional Grouping Map (United Nations Statistics Division, 2017).
27 In this case we referred mostly to documents named “national education plans”, “national strategies” or similar, usually published by the Ministry of Education of each country. In the final sample of countries, the years covered vary between 2004 and 2017.
28 Given time and resource restrictions, we were unable to include plans in other languages.
29 The template was completed with the support of the UIS.
30 For a full list of the indicators to monitor the progress toward SDG 4 on education, see the global and thematic indicator frameworks in UNESCO (2016b) or visit the UIS website (http://uis.unesco.org/en/topic/sustainable-development-goal-4).
31 In particular, from the list of 6 thematic indicators related to participation (and completion) in SDG Targets 4.1 and 4.2, those usually found in national education plans are the following: gross intake ratio to the last grade (primary, lower secondary education), completion rate (primary, lower secondary, upper secondary education), and gross early childhood education enrolment ratio (pre-primary education, early childhood educational development). However, the following indicators are rarely found: percentage of children over-age for grade (primary, lower secondary education), out-of-school rate (primary, lower secondary, upper secondary education), and participation rate in organized learning (one year before the official primary entry age).

Table 4.1 Indicators included in national education plans

	Among global and thematic indicators for SDG 4?	Equity?	Dimension			
			Sex	Location	Wealth	Disability
Learning						
Percentage of students proficient in reading for an education level						
Percentage of students proficient in mathematics for an education level						
Percentage of students achievieng minimum grade/score in national examinations	Yes	Yes	–	0	0	0
Percentage of students passing national examinations at the end of each education level		Yes	–	0	0	0
Percentage of studens achieving ninimum proficiency level (reading and mathematics) at end of the grade/education level	Yes	Yes	–	0	0	0
Learning achievement rates in examinations improved to “x %” by year “y”		Yes	–	0	0	0
Average percentage of correct answers in national examinations (reading and mathematics)						
Access: Participation						
Number of new entrants per education level		Yes	+	–	0	–
Gross enrolment ratio	Yes	Yes	+	–	0	–
Net enrolment rate		Yes	+	–	0	–
Gender parity index	Yes	Yes	–	0	0	0
Access: Completion						
Retention rate per grade or education level*						
Drop out rate by grade/education level		Yes	–	0	0	0
Survival rate by education level						
Completion rate by education level	Yes	Yes	++	0	0	0
Transition rate to next education level		Yes	–	–	0	0
Free Education						
Education provision per education level guaranteed for all by year “x”						
Abolish fees for education level “x” by year “y”						
Number of years of compulsory schooling	Yes, under the category “participation”					

Notes: ++ High frequency indicators: indicators most commonly found in the different education levels when compared to other indicators for the same category (i.e. learning, access, free education) in the reviewed national education plans.
+ Frequent indicators: indicators commonly found in the different education levels when compared to other indicators for the same category (i.e. learning, access, free education) in the reviewed national education plans.
- Low frequency indicators: indicators least commonly found in the different education levels when compared to other indicators for the same category (i.e. learning, access, free education) in the reviewed national education plans.
0 Indicators: indicators not found or almost not found in the different education levels when compared to other indicators for the same category (i.e. learning, access, free education) in the reviewed national education plans.

Source: Authors' analysis of 75 national education plans.

Some of the indicators for participation included in national plans are more prevalent than others. The most common are gross enrolment ratios and net enrolment rates. Very few national plans include measures to track the number of children out of school, which would provide an indication of the scale of the challenge in absolute terms. Sierra Leone's national plan includes an indicator for the percentage of children aged 6 to 11 who have never attended school, aiming to reduce this from 23% in 2010 to 5% by 2018.[32] Uzbekistan's plan measures participation rates in alternative programmes of education for students out of school. Even so, some countries do show concern for incorporating students who are out of primary school due to, for instance, their remote location or armed conflict. Such countries usually establish specific targets for students affected by armed conflict (such as Afghanistan and Mali) or define explicit indicators for participation of students in rural areas (as in the case of Somalia and Niger).[33]

The education plans reviewed are not generally explicit about indicators for tracking progress for over-age children as recommended by the SDGs. Yet from an equity perspective, such indicators are important. Evidence shows that being above the official school age is both most prevalent among more disadvantaged children and detrimentally related to the ability of children to complete a full cycle of primary school. In some contexts, it is also related to learning deficits (Lewin and Sabates, 2012).

More generally, access indicators tend to focus on the whole population at the given education level, rather than providing dimensions of equity for tracking progress by different population groups. As such, they apply minimum standards rather than an impartiality approach to equity *(see Chapter 2).* Among the 75 countries reviewed, a large majority provide indicators related to participation (87% of countries at the primary level and 83% at the secondary level); 73% of countries include indicators for completion of primary education and 63% of secondary education.

In education plans, participation and completion indicators feature more heavily than learning indicators *(see **Figure 4.1**).* Around 50% of these include indicators with an equity dimension for participation at each level, with a much smaller proportion doing so for completion (27% for primary and 21% for secondary education).

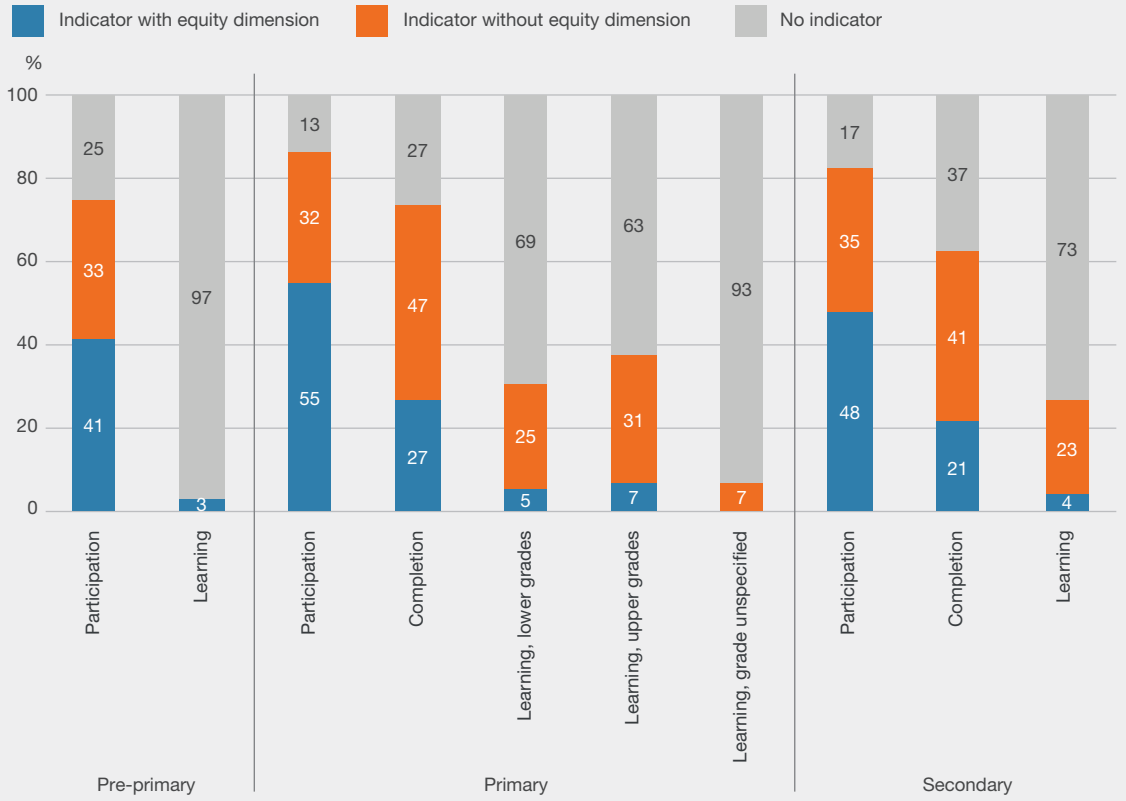
The Somalian case is especially interesting with respect to tracking progress according to equity. This country's national education plan includes measures for disaggregating enrolment information by sex, disability, special needs and location. Somalia has measured some of these indicators in the past, and thus has a baseline, which allows it to produce more accurate targets, for example:[34]

- In general, for access to primary education: "By 2016, at least 75% of children aged between 6 and 13 years are enrolled in basic primary education."[35] (p. 25)
- By student location: "By 2016, the nationwide rural-urban disparity in primary enrolment will be less than 20%."[36] (p. 25)
- By student type of community: "Increase the participation of children of pastoralist communities from about 10% to at least 40% by 2015." (p. 34)
- By sex: "Increase the enrolment of girls from the present estimate of 38% to 50% of the total primary school population by 2015." (p. 34)

32 Ministry of Education, Science and Technology, Government of Sierra Leone (2013). Education Sector Plan 2014-2018. Learning to Succeed (p. 23).
33 For Somalia, only data from the Somaliland plan were included in the analysis.
34 See Ministry of Education and Higher Education, Republic of Somaliland (2013). Somaliland's Education Sector Strategic Plan 2012-2016.
35 "Basic primary education" refers to primary education in Somaliland's education system.
36 The data set used to measure net enrolment rates in 2011 were the Primary School Census 2011/2 results, which was conducted across 13 regions in Somaliland in October 2011, with the support of UNICEF.

Figure 4.1 In education plans, participation and completion indicators feature more heavily than learning indicators

Proportion of countries with indicators for access and learning by school level and availability of equity dimensions (out of 75 countries)



Source: Authors' review of 75 national education plans.

- By student disability: "Improve the participation of male and female children (aged 6-17 years) with disabilities and those in need of special care and protection (particularly girls) to more than 40% of their share of the population by 2015."[37] (p. 34)

In most countries, student wealth or socioeconomic status does not seem to be widely used for tracking progress for access. For instance, in primary education, only Rwanda and Zimbabwe include such indicators. In Zimbabwe's case, the attendance rate of the poorest quintile in primary education is expected to rise to 94% by 2019. The country uses data from

37 This indicator refers to children in school age (6 to 17 years in Somalia). Target groups of learners with special education needs include learners with hearing impairments; visual impairments; physical disabilities; cerebral palsy; epilepsy; mental disabilities; Down syndrome; autism; behaviour, emotional and social difficulties; specific learning difficulties/learning difficulties; speech and language difficulties; multiple disabilities; chronic health problems; learners who are gifted and talented; and learners who are deaf and blind.

the Multiple Indicator Cluster Survey[38] that had a baseline of 90.4% in 2014.[39]

In the case of Rwanda, the plan refers to the EICV3[40] household survey showing that, among primary school-aged children out of school, most are from poorer households or living in rural areas. According to EICV3 data, primary education attendance rates are 9% higher in the richest consumption quintile compared with the lowest quintile.[41]

These examples use household surveys to track progress according to wealth. One reason why some national education plans do not include such indicators is because, where they rely on administrative data to enable them to track progress, these usually do not provide disaggregated data for these groups. Section 4.2 discusses the type of data needed for such tracking, notably the importance of linking school administrative data usually available through education management information systems (EMIS) with household survey data.

Although it is not straightforward to measure, completion of primary or secondary education cycles is commonly recognised as an important indicator for tracking progress. The two measures recommended for SDG 4.1, namely the gross intake ratio to the last grade (in primary and lower secondary education) and the completion rate (for primary, lower and upper secondary education) are used in some national education plans. Most of the countries that do include a completion rate have more recent plans, drawn up since 2011.[42] The definition of the completion rate is not usually very explicit, and often the calculation method is not identified in the plan, so it is not possible to verify whether the definition of these indicators is in accordance with that included in the SDG monitoring framework.[43]

Mozambique is one example of a country that produces a completion rate. Its 2013 plan identifies a "Gross Primary Education Completion Rate"[44] of 49% for all students and 45% for male students by 2010. The plan expects to achieve 54% and 51% respectively by 2015.[45] Nicaragua has set completion rate targets of 85% in 6th grade and 62% in 9th grade by 2015, with baseline values of 72% and 50% respectively in 2010.[46] In addition, Rwanda's 2013 national education plan aims to increase the primary completion rate from 72.7% in 2012 to 75% in 2018.[47]

38 The Zimbabwe Multiple Indicator Cluster Survey (MICS) was carried out in 2014 by the Zimbabwe National Statistics Agency (ZIMSTAT) as part of the global MICS programme, with technical and financial support by the United Nations Children's Fund (UNICEF).

39 Ministry of Primary and Secondary Education, Republic of Zimbabwe (2015). Education Sector Strategic Plan 2016-2020 (p. 52).

40 The third Integrated Household Living Conditions Survey or Enquête Intégrale sur les Conditions de Vie des ménages (EICV3).

41 Republic of Rwanda (2013). Education Sector Strategic Plan for 2013-2018 (p. 24).

42 Countries with a completion rate include Bangladesh, Botswana, Brazil, Cambodia, Congo, Eritrea, Gambia, Georgia, Ghana, Guinea-Bissau, Haiti, Honduras, Kenya, Lesotho, Liberia, Mali, Mauritania, Mozambique, Myanmar, Nicaragua, Niger, Nigeria (Kano State), Papua New Guinea, Philippines, Rwanda, Sao Tome and Principe, Sierra Leone, Solomon Islands, Somalia, South Africa, Sudan, Uganda, Zambia and Zimbabwe.

43 For instance, in the SDG monitoring framework, the "completion rate for primary education" is defined as: "Percentage of a cohort of children or young people aged 3-5 years above the intended age for the last grade of primary education who have completed that grade. The intended age for the last grade of primary education is the age at which pupils would enter the grade if they had started school at the official primary entrance age, had studied full-time and had progressed without repeating or skipping a grade."

44 According to Mozambique's national plan: "This indicator relates to the number of children completing primary education (Grade 7, daytime and evening classes, public, private and community education), irrespective of their age (nominator), with a 12-year-old population (denominator)" (p. 31).

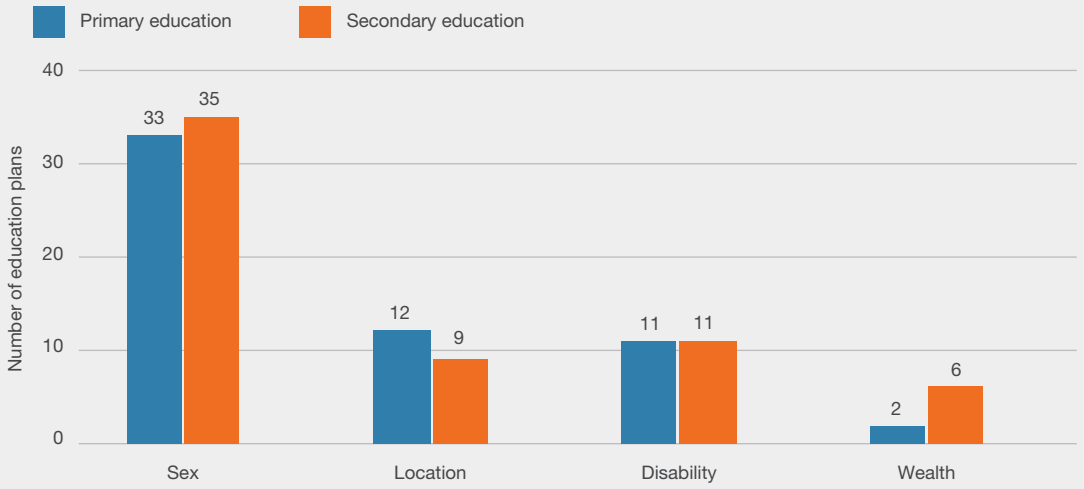
45 Ministry of Education, Republic of Mozambique (2013). Education Strategic Plan 2012-2016 (p. 136).

46 Ministerio de Educación, República de Nicaragua (2011). Plan Estratégico de Educación 2011-2015 (p. 61). Information on how the completion rate is calculated is not included in the national plan.

47 Ministry of Education, Republic of Rwanda (2013). Education Sector Strategic Plan for 2013-2018 (pp. 21 and 36). The plan does not define how the primary completion rate is calculated.

Figure 4.2 The majority of equity indicators for participation focus on sex

Number of countries including indicators for participation in primary and secondary education by equity dimension (out of 75 countries)



Source: Authors' review of 75 national education plans.

With respect to the dimensions of equity included for indicators to measure participation and completion, sex is again the most frequent *(see **Figures 4.2** and **4.3**).* Only five countries include location as an indicator of equity for completion of primary school and only two countries include in their plan regional differences for completion of secondary school. None of the education plans reviewed incorporated wealth as an equity dimension for completion of primary or secondary cycles of education.

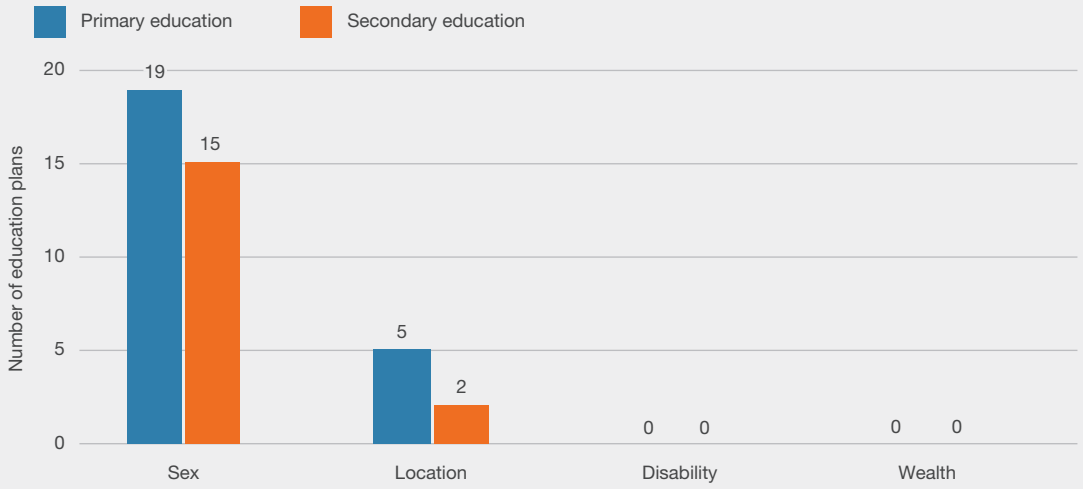
Cambodia and Sierra Leone's education plans incorporate location as a criterion for disaggregating completion rates in secondary education. Cambodia's 2014 plan stipulates that by the academic year 2017/18, 17 provinces should have achieved lower secondary completion rates of at least 40%, with a baseline of 7 provinces in 2012/13.[48] Sierra Leone's 2013 plan states that the Ministry of Education expects that by 2018, "the Completion Rate is 75% by location and gender" (from 57% for the whole population in 2011)[49], although it is not clear from the national plan what categories were defined for location and what definition of completion rate was used.

Guinea is an example of a country that intends to track progress in primary school completion by both location and sex. The country's 2014 plan identifies the starting point and target for the following indicators:[50]

48 Ministry of Education, Sport. Kingdom of Cambodia (2014). Education Strategic Plan 2014-2018 (p. 13).
49 Ministry of Education, Science and Technology, Government of Sierra Leone (2013). Education Sector Plan 2014-2018. Learning to Succeed.
50 See Tableau 24: Indicateurs de résultats du PSE in Republic of Guinea (2014). Programme Sectoriel de l'Education 2015-2017 (p. 87).

Figure 4.3 Among equity indicators on completion, sex dominates

Number of countries including indicators for completion by equity dimension (out of 75 countries)



Source: Authors' review of 75 national education plans.

- Completion rate[51] in primary education (all students): 70.7% by 2017 (baseline 58.5% in 2013).
- Completion rate in primary education (female students): 62.2% by 2017 (baseline 50.9% in 2013).
- Completion rate in primary education (students in rural areas): 50.5% by 2017 (baseline 42.3% in 2013).

Fewer education plans incorporate indicators for access to pre-primary education than for access to primary and secondary education. Three-quarters of the 75 countries include indicators for access at the pre-primary level. However, only 41% include equity dimensions *(see Figure 4.1)*. This is despite evidence that inequalities will widen over the lifecycle, unless they are tackled even before children start primary school (Alcott and Rose, 2017).

Overall, indicators measuring equity with respect to participation and completion are usually limited in national education plans. Where these are included, they most frequently relate to sex and regional dimensions and, in just a few cases, to disability or wealth. The equity indicators in national education plans tend to relate to achieving minimum standards. In the few cases where a particular population has been singled out for improving equity in educational access, indicators usually refer to the absolute and relative targets for such populations.

Parity indices, associated with impartiality measures of equity as identified in Chapter 2, are fairly common in national education plans. While these can be useful for identifying whether population groups have access to the same opportunities, parity indices have limitations as they do not show whether the overall level reached for all population groups is sufficient. For example, parity can be reached by the rich and poor

51 The definition of completion rate is not detailed in Guinea's education plan.

segments of the population even though both may have extremely low participation rates.

As an example, Sudan's 2012 education plan identified a goal of 0.95 for the gender parity index for primary education by 2016, with a baseline of 0.9 in 2010.[52] However, the primary gross enrolment ratio for 2010 was 71%, suggesting gender parity could be achieved while a large number of both boys and girls are not in school.

Uganda's 2013 education plan reports the country had achieved a gender parity index of 0.99 by 2008, which means that there is gender parity in access at the national level. However, according to the plan, national averages mask regional disparities and in many areas girls' participation is still low. Also, the completion rate of primary school for girls in Uganda is lower than for boys, and fewer girls than boys sat examinations at the end of primary school.[53]

Learning

Raising learning levels equitably at all education levels is one of the central objectives of the SDG agenda. Yet there is a striking lack of indicators for learning stipulated in national education plans, and in cases where indicators for learning are included, there is little reference to equity. At the primary level, where education indicators are most prevalent, only 37% of the 75 countries have an indicator for learning in the upper grades of the primary school cycle and just 7% of these countries have indicators for equity in learning at this *level (see Figure 4.1)*. Three additional countries – Honduras, Nicaragua and Somalia – have learning indicators for primary schooling but do not specify the grade to which their indicators pertain. The lack of learning targets corroborates findings from similar analysis conducted for the EFA Global Monitoring Report and the Global Education Monitoring Report (UNESCO, 2012; UNESCO, 2016a).

It could be argued that the lack of learning indicators in national education plans stems from the neglect of learning in the MDGs, which only focused on primary school completion. Countries with recent education plans are starting to include learning indicators. However, this is not consistent across all countries that have recently prepared their education plans. Of the eight plans included in the analysis that were produced since 2015, five include a learning indicator, only two of which are disaggregated by sex, as the only equity dimension.[54]

4.1.3 At what stages in the education cycle is equity in learning being measured?

Most countries refer to their national examinations as the basis for tracking progress in learning. This is usually done at the end of the primary or secondary school cycles. In some countries, however, national examinations are used to track progress in learning at earlier grades of primary school. For instance, Lesotho, Mexico and South Africa have one standardised examination at the end of the 3rd grade and a second one in the 6th grade, which coincides with the end of primary school.

El Salvador, Ethiopia, Honduras, Niger and Rwanda use national assessments in more than one grade of primary education. Ethiopia's 2015 national plan, for example, tracks progress on standardised examinations in different grades. The national plan measures learning achievement and sets equity targets by disaggregating measures for males and females for the indicator "percentage of students who achieve 50% and above (composite score) in the National Learning Assessment (NLA)", for students in Grades 4, 8, 10 and 12 separately. Ethiopia's plan

52 Ministry of General Education, Republic of Sudan (2012) p. 37.
53 Ministry of Education and Sports, Government of Uganda (2013). Education Sector Plan 2010-2015, p. 26.
54 Ethiopia and Mali.

identifies that the baseline measure for the Grade 4 learning indicator was 25% in 2012, which meant that 25% of students in 2012 achieved 50% and above in the composite score in the NLA. Including this baseline indicator is important to enable the government to set realistic targets for improving learning over time. The Ethiopian national education plan additionally includes the target of 50% of boys and 50% of girls achieving 50% and above in the NLA in Grades 4 and 8 by 2017, and 50% of boys and 50% of girls achieving 50% and above in the NLA in Grades 10 and 12 by 2019.[55]

The analysis of national education plans shows that 28 of the 75 countries include indicators for learning at the end of primary school *(see Figure 4.1).* A total of 23 education plans include indicators for learning at earlier grades of primary school. As mentioned previously, this is important because tracking progress from an early stage is key to ensuring that inequality gaps are tackled before it becomes more difficult and costly to do so (Rose and Alcott, 2016). But while some education plans include indicators for learning at early stages of primary school and at the end of the primary school cycle, equity in learning is mostly neglected. Of the 28 education plans with indicators for learning at the end of primary school, only 5 include disaggregation by equity dimensions.[56] Of the 23 education plans with indicators for learning in the early grades of primary school, only 4 include equity dimensions.[57]

National education plans that include targets for learning contain several types of indicators (*see* ***Tables 4.2*** *and* ***4.3***). The most common indicators are average achievement rates and percentage of students reaching a certain score or level of proficiency in national assessments, mostly measuring literacy and numeracy. In 13 of the 75 national education plans reviewed, learning targets are measured by the proportion of the student population achieving a certain minimum proficiency level at the end of the first and second cycles of primary school. This is a measure of impartiality as outlined in Chapter 2, Table 2.2.

Some similarities are found between these common indicators and those included in the thematic indicator framework for SDG 4-Education 2030 (UNESCO, 2016b). Some countries use indicators for learning by measuring the proportion of students achieving a minimum level of proficiency in literacy and numeracy, similar to SDG Indicator 4.1.1.[58] However, very few countries incorporate disaggregation for these indicators by sex, location and wealth, as recommended for SDG monitoring. When any dimension of equity in learning is used, sex is by far the most common.

Despite the strong evidence on the importance of tackling inequalities in learning early on that has been noted, of the 75 countries only Cameroon and Niger include an indicator related to learning at the pre-primary level *(see Figure 4.1).* However, neither country disaggregates for tracking progress by any equity dimension at this level. In Niger's 2013 national plan,[59] the proportion of students achieving the basic competences[60] to successfully start primary education is included as an indicator, with a goal of reaching 80% by 2024.

55 Federal Ministry of Education, The Federal Democratic Republic of Ethiopia (2015). Education Sector Development Programme V (2015) (p. 127).
56 Benin, Ethiopia, Mexico, Samoa and Zimbabwe.
57 Chad, Mexico, Samoa and Zambia.
58 SDG Indicator 4.1.1 is the "proportion of children and young people (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex" (United Nations, 2017).
59 République du Niger (2013). Programme Sectoriel de l'Education et de la Formation (2014-2024), p. 69.
60 The plan does not include the competences assessed or offer more detail about this evaluation at the end of pre-primary education.

Table 4.2 Indicators of learning, by dimensions of equity, for lower grades of primary education in national education plans

Indicators for learning	Any	None	Sex	Location	Disability	Wealth
Percentage of students proficient in reading	3	2	1	0	0	0
Percentage of students proficient in mathematics	2	1	1	0	0	0
Percentage of students achieving minimum grade/score in national examinations	2	1	1	0	0	0
Percentage of students passing national examinations in the lower grades of primary education	0	0	0	0	0	0
Percentage of students achieving minimum proficiency level (on reading and mathematics) at the end of the grade/education level	13	11	2	0	0	0
Learning achievement rates in examinations improved to “x %” by year “y”	5	5	0	0	0	0
Average percentage of correct answers in national examinations (reading and mathematics)	2	2	0	0	0	0

Note: Columns show number of plans with this type of indicator (out of 75 country plans).
Source: Authors' review of 75 national education plans.

Table 4.3 Indicators of learning, by dimensions of equity, for upper grades of primary education in national education plans

Indicators for learning	Any	None	Sex	Location	Disability	Wealth
Percentage of students proficient in reading	1	1	0	0	0	0
Percentage of students proficient in mathematics	1	1	0	0	0	0
Percentage of students achieving minimum grade/score in national examinations	6	4	2	0	0	0
Percentage of students passing national examinations in the upper grades of primary education	3	1	1	1	0	0
Percentage of students achieving minimum proficiency level (on reading and mathematics) at the end of the grade/education level	11	9	2	0	0	0
Learning achievement rates in examinations improved to "x %" by year "y"	9	9	0	0	0	0
Average percentage of correct answers in national examinations (reading and mathematics)	1	1	0	0	0	0

Note: Columns show number of plans with this type of indicator (out of 75 country plans).
Source: Authors' review of 75 national education plans.

Cameroon seeks to measure the proportion of students achieving over 20 points in a future national examination, expecting 66% of its students to attain this result by 2020. In 2013, there was still no baseline measure for this test in the national plan, as the examination had not yet been implemented.[61]

At the secondary level, only 20 of the 75 countries track learning, and only Ethiopia, Mexico and Somalia include disaggregation by equity *(see **Table 4.4**)*. In these three cases, disaggregation is by sex. For the national education plans that include measures of learning in secondary school, the three main indicators are: pass rates in examinations at the end of the secondary school cycle, average achievement rates in national examinations, and percentage of students achieving a minimum level of performance or certain proficiency level in national examinations, which usually take place at the end of secondary school. This is a measure of minimum standards *(see Chapter 2).*

Most countries that include indicators for learning at the secondary level cover the subjects of mathematics and reading in the main national language or languages (in cases when two or more languages are spoken). Some go beyond this; for instance, Malaysia and Samoa include English as a foreign language, and Guyana and Niger cover science in national examinations. Ethiopia includes physics, biology and chemistry in its national learning assessment.

Some of these countries also include learning targets for these subjects (although none of them include equity dimensions). For instance, the

Table 4.4 Indicators of learning, by dimensions of equity, for secondary education in national education plans

Indicators for learning	Any	None	Sex	Location	Disability	Wealth
Percentage of students proficient in reading	0	0	0	0	0	0
Percentage of students proficient in mathematics	0	0	0	0	0	0
Percentage of students achieving minimum grade/score in national examinations	5	3	2	0	0	0
Percentage of students passing national examinations at the end of secondary education	5	5	0	0	0	0
Percentage of students achieving minimum proficiency level (on reading and mathematics) at the end of the grade/education level	3	3	0	0	0	0
Learning achievement rates in examinations improved to "x %" by year "y"	9	8	1	0	0	0
Average percentage of correct answers in national examinations (reading and mathematics)	1	1	0	0	0	0

Note: Columns show number of plans with this type of indicator (out of 75 country plans).
Source: Authors' review of 75 national education plans.

61 Ministry of Economy, Planning and Regional Development, Republic of Cameroon (2013). Document de Stratégie du Secteur de l'Education et de la Formation (2013-2020), p. 111.

Malaysian education plan includes the following target: "By the end of Form 5 (Grade 11), 90% of students will score a minimum of a Credit in SPM [the Malaysian Certificate of Education] Bahasa Malaysia, and 70% in SPM English (against Cambridge 1119 standards) by 2025."[62] In the case of Guyana, targets for the Caribbean Secondary Education Certificate are included, with the percentage pass rate of Grades 1 to 3 mathematics, English and science in public secondary schools targeted to improve to 60% by 2018.

In sub-Saharan Africa, 18 out of 34 countries included in the review have indicators of learning either for lower or upper primary school, and only 7 of these intend to track progress by any dimension of equity at these levels: Benin, Cameroon, Ethiopia, Mali, Somalia, Zambia and Zimbabwe. Of these seven education plans, three were launched in 2015 or later, and none of them before 2011.

Where dimensions of equity in learning are covered in national education plans, disaggregation by sex is by far the most common. None of the 75 countries reviewed tracks equity in learning according to students' wealth or disability status.

Although it ought to be possible for national assessments to track progress by geographical disparities in learning, Zimbabwe is the only country among those reviewed that takes location into account. It is also notable that this is a rare example of including disaggregation by two dimensions simultaneously: sex and location. Zimbabwe's 2016 national education plan commits to specific targets in learning for students from different districts. In addition to specifying that by 2020, Zimbabwe expects to have a Grade 7 pass rate of 54% for all students (53% for boys and 55% for girls),[63] the plan further stipulates that by the same year, 45 out of 72 districts will achieve a Grade 7 pass rate of 50% or more in mathematics. With respect to disaggregation by sex, 40 districts are targeted to achieve a pass rate of 50% or more in mathematics for boys and 46 are expected to do so for girls.[64]

Six of the nine countries in Latin America that were reviewed include measures of learning at the primary level,[65] although only Mexico measures equity in both lower and upper primary levels. The Mexican National Institute for Educational Assessment applies the national standardised examination (EXCALE) to a representative sample of students at the end of pre-primary education, in the 3rd and 6th grades of primary education, and in the 3rd grade of secondary education. By assessing the same cohorts of students over time, it is possible to track progress during their school years. EXCALE provides information on baseline measures of learning and progress in learning. It includes an indicator tracking the proportion of students achieving levels of learning defined as "below average" for language and mathematics, and has defined specific targets for each grade, which include disaggregation by sex. For example, for the indicator "Percentage of students achieving learning levels equivalent to "below average" in EXCALE national examination in mathematics and Spanish", the goal was a maximum of 10.1% of female students in mathematics by 2018.[66] There are also future learning targets, which are obtained using predictions from the EXCALE data.

In the eight countries from Northern Africa and Western Asia for which education plans were analysed, only Armenia and Yemen measure educational achievement using international assessments. In both cases, they refer to the Trends in International Mathematics and Science Study

62 Ministry of Education, Malaysia (2013). Education Blue Print 2013-2025, p. E12.
63 Ministry of Primary and Secondary Education, Republic of Zimbabwe (2016). Education Sector Strategic Plan 2016-2020, p. 92.
64 Ibid.
65 Namely, Brazil, El Salvador, Guyana, Honduras, Mexico and Nicaragua.
66 Secretaría de Educación Pública, Estados Unidos Mexicanos (2013). Programa Sectorial de Educación 2013-2018 (p. 82).

Box 4.1 The Global Partnership for Education's approach to measuring equity in education planning

The Global Partnership for Education (GPE)'s strategic plan (Vision 2020) emphasises the importance of developing credible education sector plans to support stronger education systems that are equipped to deal with the challenges of equity, efficiency and learning. To this end, its recently elaborated Results Framework includes a series of indicators that enable the measurement of progress towards sector goals across its developing country partners. Key among these is a methodology for assessing the proportion of education plans that include a robust equity strategy capable of responding to the particular issues faced by marginalised groups.

Baseline data to inform this GPE Results Framework indicator were collected from 19 education plans endorsed in 2015 and 2016, including national education plans from Bangladesh, Cambodia, Central African Republic, Congo, Guinea, Guyana, Haiti, Kenya, Mali, Mozambique, Rwanda, and Togo, as well as five state-level education plans from Nigeria and two from Pakistan.

Of the 19 plans reviewed, most considered equity from a broad perspective, identifying strategies for an average of six marginalised populations *(see **Table 4.5**).* The most commonly cited were children with disabilities, girls, children from the poorest households and children from rural/remote areas. GPE methodology focuses on the equity strategy needed to address whichever group is identified as being most marginalised: usually girls (47% of plans) and children from rural/remote areas (26% of plans).

Table 4.5. Frequency with which specific marginalised groups were cited in 19 education plans

	Number of plans with a strategy for group	Number of plans where group was identified as most marginalised
Children with disabilities	18	1
Girls	17	9
Children from the poorest households	16	2
Children in rural/hard-to-reach/remote areas	14	5
Orphans	8	0
Ethnic and/or linguistic minorities	7	1
Children affected by HIV/AIDS	7	0
Refugees and internally displaced people (IDPs)	5	1
Children affected by conflict and crisis	5	0
Working children	5	0
Religious minorities	4	0
Boys	3	0
Low-demand populations (e.g. pastoralists)*	3	0
Street children	2	0

Notes: The methodology also contained a category for "other". This category included children from urban slums, as well as overage children and out-of-school children broadly. In no case was an "other" group identified as being the most marginalised.
* "Low-demand populations" refers to groups who view the formal education system as failing to teach the kinds of skills they view as necessary/useful to their livelihoods (i.e. formal education is not seen as highly relevant to, or compatible with, their lifestyles).
Source: GPE analysis based on review of 19 education plans.

A key question, then, is how effective these strategies are. Ideally, a strategy should clearly specify a long-term goal, medium-term objectives and targets. A strategy should be based on verifiable evidence (typically collected from an education sector analysis), and should be presented in parallel with a clear monitoring and evaluation

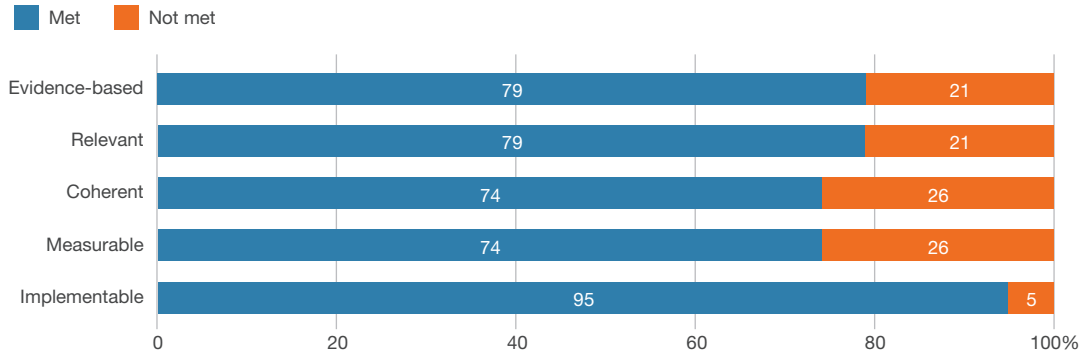
framework, including considerations of resource allocation, and roles and responsibilities. The strategy's efficacy is assessed along five key dimensions, namely whether it is:

1. Evidence-based: including identification of the underlying causes of the challenge;
2. Relevant: addressing the underlying causes of the challenge;
3. Coherent: aligning the action plan to the strategies;
4. Measurable: through the inclusion of indicators with targets; and
5. Implementable: identifying cost, funding source, responsible entity and timeframes for operationalisation.

A significant limitation of the GPE methodology stems from its review being desk-based. This means that the Results Framework indicator on the proportion of education plans that include a robust equity strategy cannot capture levels of national ownership of, and political buy-in to, the identified equity strategies. Yet these elements are crucial to effective implementation and should, in theory, be considered in parallel to a full assessment of the quality of the strategy. However, evaluating the political credibility of any planning document would require complementary, more qualitative methodological approaches, which are beyond the scope of the GPE indicator.

Results from the analysis of baseline data are presented in **Figure 4.4**. It is encouraging that all 19 plans included at least some reference to strategies for addressing equity issues, thus highlighting the importance accorded to tackling disparities in sector planning processes. In addition, performance overall is strong, with almost three-quarters of plans meeting at least four of the five dimensions. A clear focus on implementation is apparent, with only one plan failing to define how the strategy for marginalised populations would be operationalised. In five cases, measurability was assessed as poor; the extent to which this reflects the difficulty of tracking data for marginalised groups vs. a more general issue of poor capacity to develop robust monitoring frameworks warrants further investigation.

Figure 4.4 Performance of 19 education plans against GPE quality standards



Source: GPE analysis based on review of 19 education plans.

The rating was conducted on a graduated scale, with a dimension considered as being "met" when the plan reflected a reasonable effort to elaborate the core elements of that dimension. Three of the 19 plans scored the maximum possible on all five dimensions of the assessment (i.e. the plan articulated all elements of all dimensions): those for the Central African Republic (refugees and internally-displaced people (IDPs)), Guinea and Togo (both children in remote/rural areas). It is encouraging that the former two are transitional education plans, underscoring that in situations of crisis and fragility marginalised populations remain a key sector priority.

Source: Analysis prepared by the Global Partnership for Education.

(TIMSS) as the basis for their data for tracking progress. Yemen also refers to the Arab Knowledge Report Test. Yemen tracks literacy in lower primary education and mathematics and science in 6th grade. None of the plans reviewed in the region measures equity in learning.

Of the nine plans analysed in Central and Southern Asia, only Nepal, the Sindh Province in Pakistan, and Uzbekistan track learning at the end of primary school.[67] Whereas for Nepal and the Sindh Province in Pakistan literacy and numeracy are assessed in 5th and 6th grades respectively, Uzbekistan defines and assesses education standards at the end of general secondary education (Grades 5 to 9). Again, none of the countries includes a measure for equity in learning.

Regarding the five countries in Eastern and South-Eastern Asia (Cambodia, Malaysia, Mongolia, Myanmar and Timor-Leste) for which information is available, only Malaysia includes learning indicators. Malaysia's 2013 education plan has indicators for learning in both primary and secondary education.[68]

In summary, there is a lack of attention to equity in learning across all national education plans. Of the 24 national education plans reviewed for countries which are classified as low income, 9 include learning indicators in early grades of primary education. Only 3 of them include equity dimensions in their indicators: Ethiopia, Mali and Somalia, and equity is only focused on sex. Of the 34 national education plans for lower-middle-income countries, only 11 include learning indicators and only Cameroon and Zambia disaggregate by equity dimensions. Even among the education plans for the 11 countries classified as upper-middle income, only Mexico and Samoa track learning according to equity in primary education.

4.2 DATA NEEDS FOR MEASURING EQUITY[69]

The preceding analysis of national education plans and the overview of data availability in Section 3.4 make clear that much can be done to improve education indicators on equity, and especially to ensure that the most disadvantaged children are represented. A more expansive approach to how disadvantage is measured is recommended, which would include expanding the coverage of data collection and more explicit analysis of disadvantaged groups from the earliest years.

4.2.1 A more expansive approach to how disadvantage is measured

Collecting data on a broader range of dimensions of inequity

Where country education plans do disaggregate indicators, the focus is most frequently on sex. While girls have fewer educational opportunities than boys in many contexts globally (Rose, Sabates, Ilie and Alcott, 2016), which makes sex an important dimension to track, it is just one element of potential inequality. For education plans to better tackle disadvantage, data must be collected on a far broader range of characteristics: children's socioeconomic, disability, geographic, racial, ethnic and linguistic characteristics, in addition to sex. These are all dimensions for which data are, or could be, collected in most settings.

In many contexts, the most important dimension of inequity is socioeconomic status. Although average wealth levels vary greatly across countries, inequities between the poorer and richer within countries are near-ubiquitous (UNESCO, 2014). Cross-sectional data from East Africa and South Asia show that the

67 Other provinces of Pakistan were not included in the present analysis.
68 Ministry of Education, Malaysia (2013). Education Blue Print 2013-2025 (pp. E12 and E25).
69 Parts of this section draw on Rose, Sabates, Alcott and Ilie, 2016.

learning of children from less economically-advantaged households is at least one year behind that of children of the same age from more economically-advantaged households (Alcott and Rose, 2016; Jones and Schipper, 2012). In South Africa, by Grade 3 the poorest 60% are three grade levels behind the wealthiest quintile (Spaull and Kotze, 2015).

Measuring socioeconomic status is not straightforward but is possible. Some relatively simple approaches have been developed, which are compatible with more complex measures of income and expenditure. One such measure is the wealth index, which requires recording the ownership of a set of household goods, housing characteristics and access to household services in order to compute the relative wealth position of each household within the country. Wealth indices using DHS data have been found to perform as well as expenditure data in explaining variation in educational outcomes and are as useful as other relevant indicators of child health and well-being (Filmer and Pritchett, 1999; Filmer and Scott, 2012). There have also been subsequent improvements in the computation of wealth indices to make them comparable across countries and over time (Smits and Steendijk, 2015).

Another measure of socioeconomic status that requires a relatively small number of items to compute is the poverty score (Schreiner, 2010). The idea behind the poverty score is to create a "scorecard" of objective poverty indicators that are strong determinants of income poverty in the given context. A typical poverty score card includes information from household and housing characteristics (e.g. cooking fuel, type of floor), background of the household head (e.g. single parent, highest level of education), access to household services (e.g. electricity, water) and durable goods (e.g. mobile phone, assets). A poverty score requires information from consumption and expenditure surveys which are available in many countries that make it possible to compute the income poverty line and identify the most relevant items in the given context. Once this is established, information for a simple poverty card can then be collected at scale.

A potential limitation with the wealth index and poverty scorecard though is their inability to capture changes in the predictive power of specific items on income poverty over time. For example, the lack of a mobile phone is now a strong predictor of poverty in many sub-Saharan African countries, but this was not the case a decade ago.

Overall, the key message is that with a small number of simple, easy to collect items for a given measure, which are usually also available in household surveys, education authorities can enrich their information on students and link this information to educational access, progress and completion. Advances have also been made to link children's background information to learning, for example using data from the Young Lives international research project or citizen-led assessments from the People's Action for Learning (PAL) Network *(see Section 4.2.2).* In addition, most household surveys contain information that enables the computation of wealth indices and poverty scores described above.

Another key potential dimension of inequity is regional and geographic disparities within a given country. Taking rural India as an example, there are large cross-state disparities in the proportion of 10- to 13-year-olds who are in school and learning (i.e. able to perform division) at all wealth levels. Among households from the poorest quartile with equivalent levels of deprivation, the proportion of 10- to 13-year-olds who are in school and learning ranges from 7% in Gujarat to 33% in Tamil Nadu; among households from the wealthiest quartile, the same proportion ranges from 30% in Maharashtra to 76% in Manipur *(see **Figure 4.5**).* Similarly, in South Africa, children from the wealthiest provinces are six times as likely to have basic mathematics skills as children from its poorest provinces (Moloi and Chetty, 2010). Another clear divide is between urban and rural environments. In Ethiopia, urban 8-year-olds are more than five times as

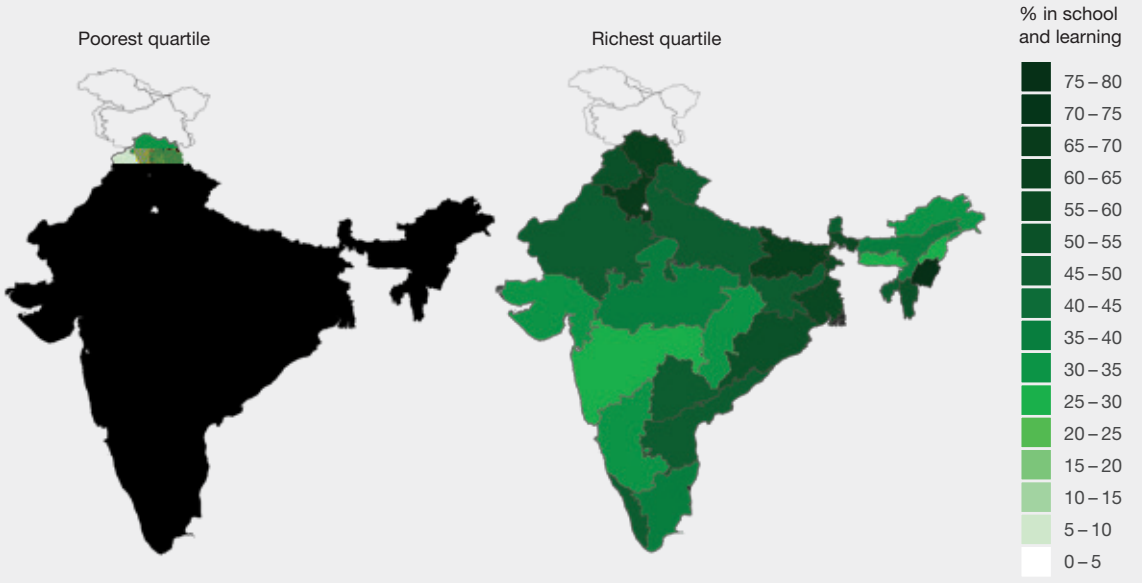
likely as rural 8-year-olds to be able to read sentences (Rolleston et al., 2014). Urban/rural inequalities are also apparent in El Salvador, Guatemala, India, Pakistan, Panama, Peru, Tanzania, Viet Nam and Zambia (Altinok, 2013; Burger, 2011; Rolleston et al., 2014; Tayyaba, 2012; UNESCO, 2014).

Collecting data on regional location is often straightforward. Most education management information systems (EMIS) datasets have information both about school resources and their geographic location. This complementarity makes it possible to have refined information about clusters of indicators, for example on the different educational backgrounds of teachers by sub-region. With national data, such as national examinations, it is also possible to obtain the location where students took their examination, which will often be a good proxy for the location where they reside. This information enables analysis of regional gaps in learning. Household surveys also collect information at the regional level, although disaggregation of indicators at sub-regional levels is not always possible. For example, DHS surveys are representative of the country and of the main regions of the country but cannot always be representative of more refined geopolitical divisions within countries. It is also possible to collect representative indicators that distinguish rural, semi-urban and urban areas, although identifying boundaries between rural and urban settings can be a significant challenge, in particular in contexts with large peri-urban populations and with rapid urbanisation.

Ethnic, racial and linguistic groupings frequently provide a further source of inequity. Children whose household language is different from their language of instruction learn less in a range of countries, including Benin, Cameroon, Guatemala, the Islamic Republic of Iran and Turkey (Altinok, 2009; Fehrler and Michaelowa, 2009). In Peru, the average

Figure 4.5 Educational opportunities vary greatly across rural India

Percentage of children in India aged 10-13 years who are in school and learning, by state and wealth, 2014



Notes: The maps displayed in the charts are for reference only. The boundaries, colors, denominations and any other information shown on these maps do not imply, on the part of the UIS, any judgment on the legal status of any territory, nor endorsement or acceptance of such boundaries. Data not shown for contested areas of Jammu and Kashmir, Shaksam Valley, and Aksai Chin. Sample covers 10- to 13-year-olds in the highest and lowest wealth quartiles. 'In school and learning' refers to the child being in school and able to perform division.
Source: Rose et al. (2016), drawing on data from ASER 2014.

primary school child whose mother tongue is Spanish outperforms 84% of children speaking an indigenous language in mathematics (Cueto et al., 2014). Collecting data on differences between these groupings (a form of horizontal inequality) is possible, but in some contexts can be politically charged. For example, it is not considered appropriate in post-genocide Rwanda, where Tutsis, Hutus and Batwas are not identified by ethnicity in learning assessments or national surveys. Nevertheless, Rwandans living in extreme poverty are sometimes identified as being "historically marginalised", although not all national surveys single out the population in this way.

Until recently, children with disabilities have been invisible in, and sometimes even excluded from, most data sets, largely due to challenges in identifying them and concerns about stigmatisation through doing so. While cultural and linguistic variations in understanding, defining and responding to disability have made this element of equity particularly difficult to measure at scale, important progress has now been made on how to remedy this. Rather than asking the very direct question in surveys, "Do you have a disabled member in your family?", international initiatives have begun to rephrase the question and to ask instead about the difficulties that children face (relative to other children of the same age).

The approach by the Washington Group on Disability Statistics towards identifying functional limitations is based on the UN Convention on the Rights of Persons with Disabilities, which defines disability as including "those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others" (United Nations, 2007). Over the years, the Washington Group has developed sets of questions aimed at different population age groups. This includes a set of "Child Functioning Questions" for children aged 5 to 17 years. To focus the respondent on the functioning of their own child with reference to that child's cohort, where appropriate, questions are prefaced with the clause "Compared with children of the same age…" Overall, experience with using the questions developed by the Washington Group indicates that they provide a simple, sensitive and nuanced way of capturing disability, even in contexts where there are concerns that stigma could prevent direct reporting of disability. The questions provide the opportunity for international comparability and have been developed using a rigorous methodology.

An example of the use of disability identifiers in a household survey comes from the Annual Status of Education Report (ASER) Pakistan which, since 2015, has used the short set of questions developed by the Washington Group with adaptions from the longer set of child functioning questions. From these data, it was possible to determine that children with disabilities are likely to be among the most disadvantaged in education. Based on ASER Pakistan data in Punjab, Pakistan, only 71% of 5- to 16-year-olds with moderate to severe difficulties were attending school, compared with the average school attendance rate of 83% for children of the same age who were not reported as having any difficulties. There is an even starker gap, though, in rates of learning: just 11% of children with moderate to severe difficulties could do subtraction, compared to 53% of children with no difficulties. Type of disability matters too: children with moderate to severe physical difficulties were twice as likely to be out of school as children with moderate to severe learning difficulties. However, none of the children with moderate to severe learning difficulties were able to do subtraction, in contrast to 15% of children with moderate to severe physical difficulties (Rose, Sabates, Alcott and Ilie, 2016).

There have been significant advances in recent times towards identifying children with disabilities in surveys in ways that allow their access to education and learning to be compared with that of other children from similar backgrounds. This has been helpful in developing new ways to provide appropriate support and resources. However, other groups remain

invisible. For example, nomadic and migratory groups continue to be hard to reach and are near absent even in population censuses (Carr-Hill, 2017). Slum populations are also hard to reach by enumerators working for national statistical agencies due to problems of security and violence, although other, citizen-led surveys have had some success reaching these populations (Carr-Hill, 2017).

At the international level, several initiatives have been launched to improve the measurement of equity in learning. International comparability of equity measures is particularly important in the SDG 4 monitoring framework, which also means that agreement on common definitions, metrics and standards is necessary. **Box 4.2** describes activities by the UIS and other international organizations in this area.

Box 4.2 International initiatives in support of equity measurement

The UIS, the statistical agency responsible for compiling and disseminating internationally-comparable data in UNESCO's fields of competence, plays a critical role in the Education 2030 Agenda. This mandate is set out in the Education 2030 Framework for Action, which ratifies that the UIS is the official source of cross-nationally-comparable data on education for SDG 4 (UNESCO, 2016b). A particularly critical aspect for reporting on progress towards SDG 4 is the measurement of equity across all data sources for education indicators. In support of this goal, the UIS has convened expert groups to work on indicator development, published methodological guidelines and expanded its international database with education indicators to improve coverage across countries, years and dimensions of disaggregation.

In 2016, the UIS convened the Technical Cooperation Group on the Indicators for SDG 4–Education 2030 (TCG) (http://tcg.uis.unesco.org) as a platform to discuss and develop the indicators used to monitor SDG 4 in an open, inclusive and transparent manner. The TCG works in tandem with the Global Alliance to Monitor Learning (GAML) (http://gaml.uis.unesco.org) to make recommendations on indicator development. GAML takes the lead in the development of indicators related to learning and skills assessment, while the TCG focuses on issues related to the other indicators. The UIS also leads two other expert groups: the Inter-Agency Group on Education Inequality Indicators (IAG-EII) (http://iag.uis.unesco.org) and one Task Force of the Intersecretariat Working Group on Household Surveys (ISWGHS) (https://unstats.un.org/iswghs/).

Household surveys are the main source of disaggregated education data for the analysis of disparities between different population groups. The IAG-EII, created by the UIS, UNICEF and the World Bank in 2016, aims to promote and coordinate the use of household survey data for SDG 4 monitoring at the national, regional and global levels. The IAG makes recommendations to harmonise the processing of survey data by different agencies and collaborates on standardised definitions of survey-based indicators and of individual and household characteristics for data disaggregation.

In the ISWGHS, the UIS leads the Task Force on Standards for Education Spending Estimates based on Household Survey Data, which focuses on methodological development of SDG Indicator 4.5.4 (education expenditure per student by level of education and source of funding). The Task Force has produced a document mapping sources of information on education spending by households (UIS, 2017) and is also drafting a *Guidebook on Education Expenditure in Household Surveys* that will be published in 2018.

In addition to these examples of methodological work, the UIS continues to improve the availability of data for monitoring of progress towards SDG 4. Recent activities include the launch of pages dedicated to SDG 4 and equity in the UIS website (http://uis.unesco.org), addition of more disaggregated education indicators in the UIS. Stat online database (http://data.uis.unesco.org) and the dissemination of interactive maps and charts that allow users to explore data related to SDG 4.

Source: Prepared by the UNESCO Institute for Statistics.

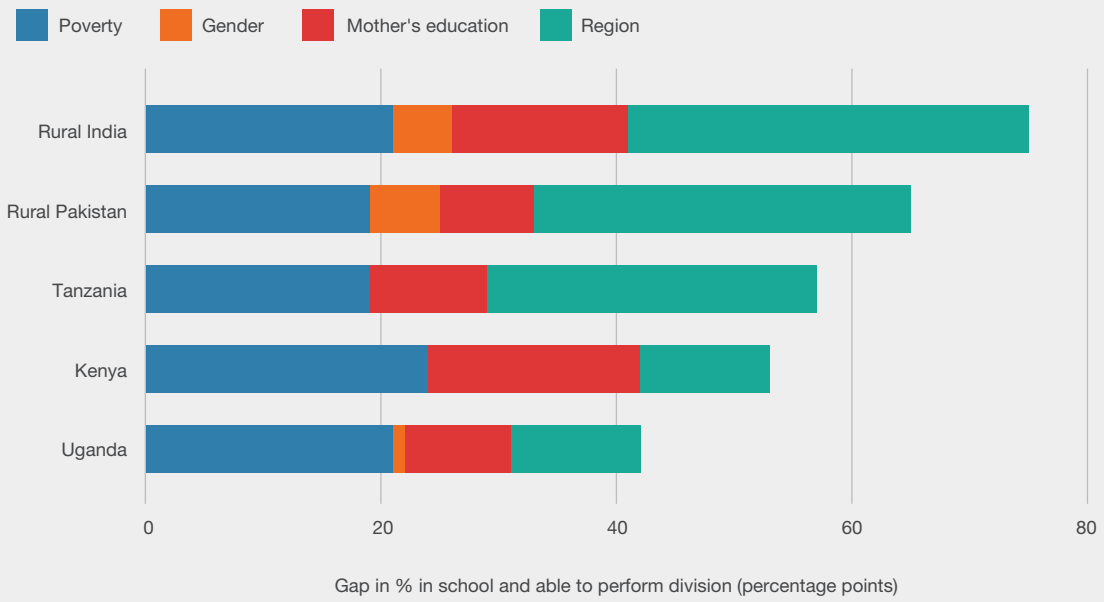
The intersecting nature of disadvantage

Sources of inequity frequently compound one another. It is therefore crucial to view child characteristics in conjunction with each other rather than in isolation. Taking sex and socioeconomic status as a starting point, it is most often poor girls who are least likely to be learning the basics when these factors interact. In rural India, gender disparities are considerable between poorer girls and boys at the primary level, while wealthier girls keep up with or exceed learning among wealthier boys (Alcott and Rose, 2017).

Across South Asia and East Africa, there are sizeable gaps between poorer and wealthier children in enrolment and learning (Rose, Sabates, Alcott, and Ilie, 2016). In Kenya, Tanzania and Uganda, and in rural areas of India and Pakistan, there is at least a 20-percentage-point gap between rich and poor in the share of children aged 10 to 13 years who are in school and have learned basic mathemathics skills *(see **Figure 4.6**)*. These gaps are considerable in and of themselves.

Figure 4.6 Socioeconomic inequities are exacerbated by other disadvantages

Percentage point gap in ability to perform division between more- and less-advantaged groups, showing cumulative impact of additional dimensions of disadvantage



Notes:
1. Sample covers 10- to 13-year-olds.
2. 'Poverty' differentiates between the highest and lowest wealth quartiles. 'Mother's education' differentiates between those whose mothers attended school and those whose mothers did not. 'Region' differentiates between locations depending on the country: state (India), province (Pakistan), region (Tanzania), county (Kenya) and sub-region (Uganda).
3. Figure 4.6 uses logistic regression models with interaction terms that progressively added more intersections of inequality. The bars build upon one other to show the increase in inequality as one moves from more general subgroup comparisons, such as between poor and rich, to more specific comparisons, such as between girls from low wealth quartiles and boys from high wealth quartiles. For each country, the first bar (blue) shows the gap between the poorest and richest quartiles, and the second bar (orange) alone shows the additional inequality between poor girls and rich boys. The total size of the poor girl-rich boy gap is the two bars added together. For a given country, when one adds all four bars together this shows the gap between (1) boys from the highest wealth quartile in the best-performing region whose mothers attended school, and (2) girls from the lowest wealth quartile in the worst-performing region whose mothers did not attend school.
Source: Rose et. al. (2016), based on the 2014 ASER and UWEZO surveys.

Yet, when taking account of multiple dimensions of disadvantage, inequalities grow further still. In rural India, rural Pakistan and Uganda, wealth gaps are compounded by gender disparities. In rural Pakistan, for example, the gap between poorer and wealthier children increases by one-third, from 19 to 25 points, when comparing poorer girls to wealthier boys. And while the occurrence of gender disparities varies by country, first-generation school-goers are at a disadvantage in all countries. When focusing not only on poor girls but on those whose mother never attended school, the gap between these children and wealthier boys whose mother did attend school increases learning inequalities in each country by at least 8 points. In Kenya, this combination of factors almost doubles the gap, from 24 to 42 points *(see Figure 4.6).* Within each country, regional disparities further exacerbate inequality, most starkly in rural India, rural Pakistan and Tanzania. In Tanzania, regional disparities double the inequality in rates of children in school and learning: the gap between wealthier boys whose mothers went to school and poorer girls whose mothers did not stands at 29 points, but this gap increases to 57 points when comparing advantaged boys in the best performing region to disadvantaged girls in the worst performing region.

The inter-sectional nature of disadvantage cannot be overlooked if policies are to support those most likely to be left behind. To improve the accuracy, relevance and efficacy of policy and planning, it is essential to collect data and track progress on multiple forms of disadvantage. This would necessarily include gender, but would also go well beyond it in order to assess how the interaction of gender with other key sources of disadvantage holds children back. As indicated in Chapter 3, this is an issue of sampling and design, as well as of the type of data collected.

4.2.2 Expanding the coverage of data collection

Data should focus on, and account for, a broader range of disadvantages. What then are the additional implications for the process of collecting these data?

Learning assessments must reach those out of school

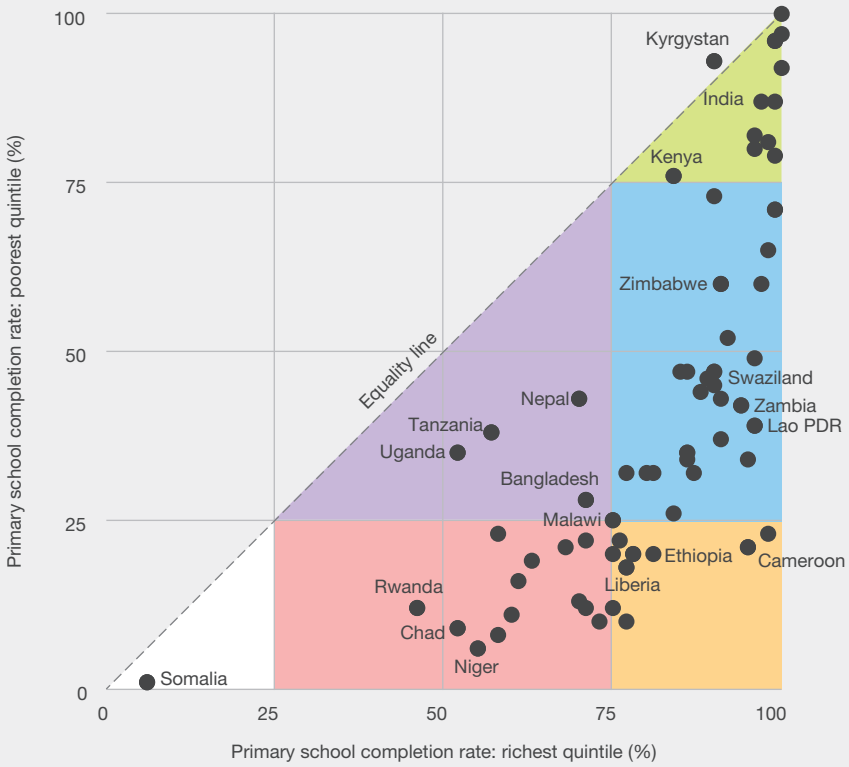
It is insufficient to assess learning of school-attending children alone. Measuring progress towards education targets means also gathering data on and including the most disadvantaged children, who are frequently not in school. Among 67 low- and middle-income countries with data, very few have achieved equality in primary school completion between the rich and poor *(see* ***Figure 4.7***) and across these countries, the average gap between the richest and the poorest is 32 percentage points. As such, using data on school-going children alone to track progress on learning would disproportionately represent the relatively well off, thus bias an understanding of conditions for the most disadvantaged.

In addition, it is not sufficient simply to presume that the disadvantaged children who are in school are representative of disadvantaged children out of school. For example, among the poorest girls in rural Pakistan who are in school, only around one-half of these children have learned basic mathematics skills by age 12 *(see* ***Figure 4.8***). While this in itself is alarming, among all poor girls who are out of school at age 12, less than 5 in 100 have gained these skills.

Complementary sources of data are therefore needed. Data sampling representative of all children might thus be better achieved by adopting a sample-based household survey of the kind used in PAL Network citizen-led assessment surveys, DHS and other household surveys, which also have the benefits of enabling the collection of data related to the background characteristics of the children and their households. PAL Network surveys, for

Figure 4.7 In almost every country, poorer children are far less likely than richer children to complete primary school, latest year available (2006-2014)

Percentage of poorest and richest quintiles who complete primary schooling



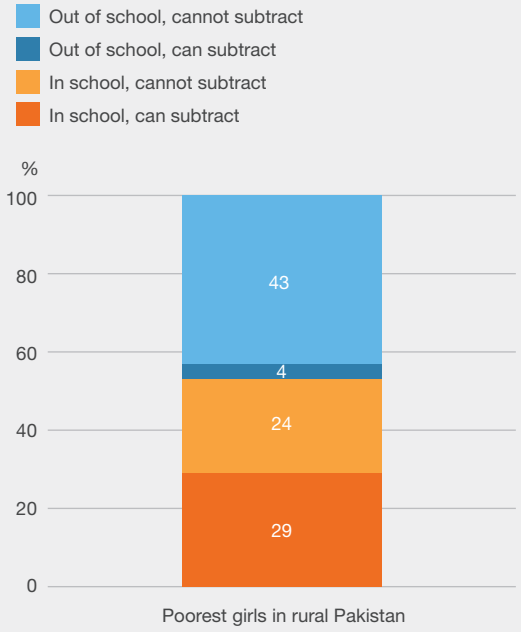
Notes:
Group 1 (green, 13 countries): High completion rates (over 75%) for both richest and poorest: Armenia, Bolivia, Egypt, Georgia, Guyana, India, Indonesia, Kenya, Kyrgyzstan, Republic of Moldova, Tajikistan, Ukraine, Viet Nam.
Group 2 (blue, 26 countries): High completion rates (over 75%) for richest and moderate completion rates (25% to 75%) for the poorest: Benin, Bhutan, Cambodia, Comoros, Congo, Democratic Republic of the Congo, Djibouti, Gambia, Ghana, Guatemala, Honduras, Lao People's Democratic Republic, Lesotho, Morocco, Nicaragua, Pakistan, Philippines, Sao Tome and Principe, Sierra Leone, Sudan, Swaziland, Timor-Leste, Togo, Uzbekistan, Zambia, Zimbabwe.
Group 3 (orange, 10 countries): High completion rates (over 75%) for richest and very low completion rates (below 25%) for the poorest: Cameroon, Ethiopia, Guinea, Haiti, Liberia, Madagascar, Malawi, Mauritania, Nigeria, South Sudan.
Group 4 (purple, 4 countries): Moderate completion rates (25% to 75%) for both richest and poorest: Bangladesh, Nepal, Tanzania, Uganda.
Group 5 (red, 13 countries): Moderate completion rates (25% to 75%) for the richest and very low completion rates (below 25%) for the poorest: Afghanistan, Burkina Faso, Burundi, Central African Republic, Chad, Côte d'Ivoire, Guinea-Bissau, Mali, Mozambique, Niger, Rwanda, Senegal, Yemen.
Group 6 (white, one country): Very low completion rates (below 25%) for both the richest and poorest: Somalia.
Source: Rose, Sabates, Alcott and Ilie (2016), using WIDE database, DHS, MICS, ASER, and Uwezo data.

example, randomly select villages and households within districts, and provide survey weights to account for the different sizes of different districts. The robust nature of this sampling approach provides nationally representative data without a need to visit all households, as in a census. For more information on the design of household survey samples, refer to Section 3.5.

Including learning assessments as part of household surveys is key to making sure that out-of-school children are included for tracking progress in learning. This also offers the additional benefit of avoiding unintended consequences of school-based assessments, such as schools putting forward their most able children and "teaching to the test". Such assessments have to be well designed to capture

Figure 4.8 Among poor girls in rural Pakistan, those out of school are far less likely to be learning

Percentage of poorest quartile of girls in rural Pakistan who can and cannot subtract, by schooling status, 2014



Note: Sample covers 10- to 13-year-old girls in rural areas who are in the lowest wealth quartile.
Source: ASER Pakistan 2014.

children's skills and competencies. To aid analysis, they should be comparable over time and of broad enough scope to capture variation across different ages. Assessments should also avoid strong floor and ceiling effects, defined by the inability of children to respond to any question in the examination or for most children to respond to all questions correctly. This is a common pitfall with national examinations in developing countries based on the competencies of the national curricula, which are frequently unrealistically difficult for children (Pritchett and Beatty, 2015), thus creating floor effects preventing any meaningful analysis of the stage most children are at in their learning.

Another shortcoming of high-stakes national examinations is that there is no comparability over time, as examinations are used to select students to pass to the next level, and so they are standardised differently every year. Changes in content and various features of the questions alter the likelihood of children answering correctly. The scores obtained by "equivalent" children in different years will therefore vary (Goldstein, 1983; Newton, 1997). Since it cannot be assumed that national examinations are comparable over time, standardisation is possible by setting cut scores for each level which reflect the equivalent achievement of the previous year (or the previous time when the national examination took place).

While household-based learning assessments improve upon school assessments in their coverage of the most disadvantaged groups, there are still important gaps. The most pressing is how to reach children not living in formal households, such as children in nomadic communities, institutions, unrecognised urban settlements, or those displaced by conflict. When data are collected by local citizens, their contextual knowledge may make such surveys particularly apt for gathering data for such children (Carr-Hill, 2017). Still, ongoing sensitivity and consideration should always be given to how data can more comprehensively represent such children.

Choices need to be made with respect to whether assessments should measure learning the basics or the competencies of a specific curriculum. Whichever choice is made, it is crucial that every effort be made to reach the most marginalised and record children's key background characteristics. Household surveys offer a good way to collect information about children that cannot easily be gathered at the school level (such as socioeconomic status) and would also include children who are not in school.

Ideally, school and household data should be linked, so as to highlight the effects of class size, facilities, teacher preparation and teaching practices on the most disadvantaged children. This could be done using administrative data already collected

on schools, such as EMIS data. The Young Lives international research project and APRESt (Muralidharan and Zieleniak, 2013) offer two current examples of how school and household data have been linked in low- and lower-middle income countries.[70] Data from Young Lives, for example, have been used to show the links between schooling opportunities and learning outcomes between more and less advantaged children (Singh, 2014).

4.2.3 Focusing explicitly on disadvantaged sub-groups from the earliest years

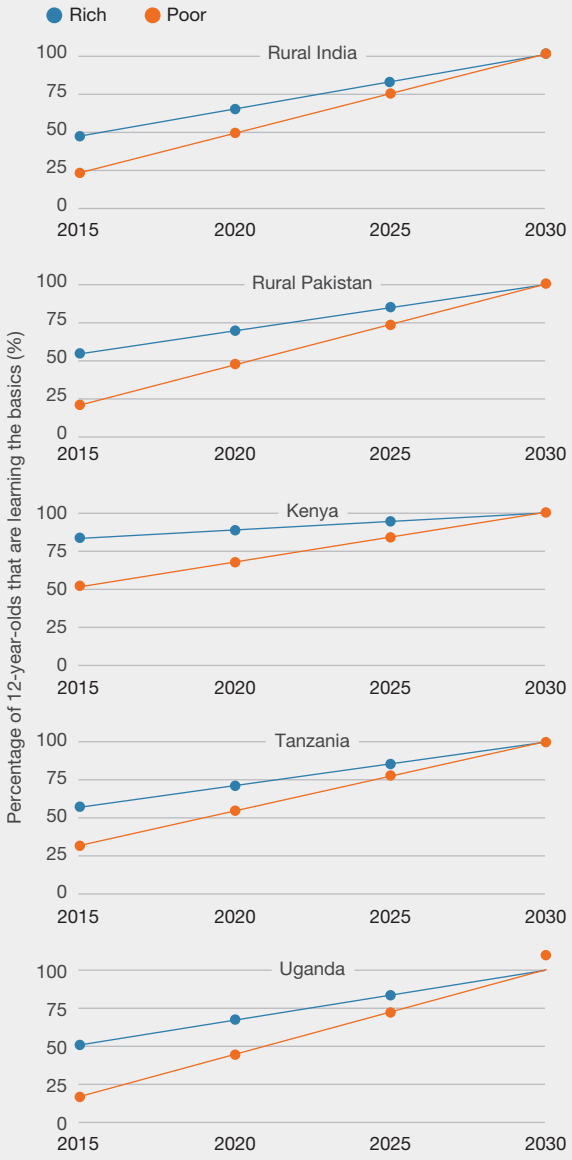
The key principles then in collecting data are to ensure identification of the most disadvantaged groups to ensure that coverage includes all children (including those not in school) and to complement these data with information on the educational opportunities available to each child. But how to make best use of such data?

Set "stepping stone" interim targets to track progress before it is too late

The Education 2030 Framework for Action accompanying the SDGs acknowledges that "no education target should be considered met unless it is met by all" (UNESCO, 2016b) and so it is crucial that data analysis disaggregates and focuses on the most disadvantaged. It is therefore key to maintain a similar focus on indicators and, by extension, policy responses. For example, in order to identify the progress needed by 2030, Watkins' (2015) proposal of "stepping stone" targets, which set interim targets adjusted to specific countries, offers a clear means to account for the different rates of progress needed for different sub-groups within countries *(see **Figure 4.9**)*. In addition, the use of stepping stone targets with shorter intervals (e.g. every five years) is more informative for the reality of national planning imperatives, which tend to relate to political electoral

Figure 4.9 Ensuring all 12-year-olds are learning the basics by 2030 will require efforts targeted at the most disadvantaged

Projected change for all children to be learning by 2030, by wealth quintile



Notes: These projections are based on ASER India 2012, ASER Pakistan 2012, and Uwezo 2012. 'Rich' and 'Poor' refer to the richest and poorest quintile of households in each country, respectively.
Source: Rose and Alcott (2015).

70 The Andhra Pradesh Randomized Evaluation Studies (APRESt) hold data from a range of policy evaluations. These data are not currently publicly available.

cycles; such targets also make it possible to see whether sufficient progress is being made for the most disadvantaged groups well before the more distant deadline year (Rose, 2015).

Track progress in the early years

In order to be able to track progress towards learning goals for all children, it is essential that data analysis actively inform practice, rather than simply describe the state of affairs once it is too late. Rather than school-leaving data (or, in some countries, secondary school entrance exam information), a greater emphasis should be put on learning assessments at earlier ages. This is when inequities begin and after which they become entrenched: in all four of the Young Lives countries – Ethiopia, India (Andhra Pradesh), Peru and Viet Nam – the richest quartile makes more progress than the poorest quartile in mathematics between ages 5 and 8 (Rolleston et al., 2014).[71] Identifying those most in need is critical to maintaining their progress: data from a range of learning assessments in India indicate that only between 9% and 13% of those who lack basic literacy or numeracy skills are able to gain this skill even after an additional year of schooling (Bhattacharjea et al., 2011; Educational Initiatives, 2010; Pritchett and Beatty, 2015).

In short, those facing the greatest educational inequities live in households affected by poverty, with disadvantage reinforced by where children live, their gender and whether they have a disability. They are also likely to be children not living in formal households and who are therefore usually excluded from data collected with household surveys. Such factors should not determine a child's learning potential. Given uneven progress through primary schooling, and that this progress is often linked to sources of inequality associated with inherited disadvantage, there is a need to track progress in learning from the earliest years. In order to better measure progress towards more equitable education systems, data must identify multiple sources of disadvantage, ensure coverage of the most marginalised populations and be put to use in time to not only describe the opportunities afforded these children but also to influence the design of policies aimed at achieving equity in education.

4.3 CONCLUSION

Ensuring no one is left behind in educational access and learning is a major priority in SDG 4. As many countries do not have regular standardised national assessments, it is difficult to establish comparisons both within and between countries. The SDGs certainly highlight the urgent need to develop assessments that can be used to compare progress towards basic literacy and numeracy. But even if they are developed, they might not be the most appropriate way to track learning from an equity perspective: notably, children who are out of school are likely to be among the most disadvantaged and remain invisible in these assessments. In addition, public examination results do not always provide comparable data in a form needed to track progress over time, nor do they enable disaggregation by core dimensions of inequality, such as socioeconomic status.

Currently, national education plans might implicitly measure equity by ensuring minimum standards (as defined in Chapter 2) are reached, i.e. that all children in principle, regardless of their backgrounds, need to achieve the same level in learning. However, unless progress is tracked for different sub-groups, it is unlikely that disadvantaged groups can improve at the faster pace they need to close learning gaps with the rest of the population.

71 Related analysis finds that the gap in achievement between these four countries remains stable as children get older, with children in Viet Nam performing best, followed by Peru, India and then Ethiopia. The gap remains wide, even once socioeconomic status and other factors are taken into account (Singh, 2014).

It is necessary to measure equity in learning at all education levels and for the whole school-age population (whether in school or not), starting from early grades. The use of more appropriate indicators to track equity in access and learning, such as those related to impartiality by disaggregating for sub-groups of the population as described in Chapters 2 and 3, will allow us to ensure that not only minimum standards are met but also that equity gaps in education are narrowed.

REFERENCES

Alcott, B. and P. Rose (2016). "Does private schooling narrow wealth inequalities in learning outcomes? Evidence from East Africa". *Oxford Review of Education.* Vol. 42, pp. 495-510. https://doi.org/10.1080/03054985.2016.1215611

—— (2017). "Learning in India's primary schools: How do disparities widen across the grades?" *International Journal of Educational Development.* Vol. 56, pp. 42-54. https://doi.org/10.1016/j.ijedudev.2017.05.002

Altinok, N. (2013). "Performance differences between subpopulations in TIMSS, PIRLS, SACMEQ and PASEC". Background paper for the *EFA Global Monitoring Report 2013/4.*

—— (2009). "An empirical approach to marginalization in education based on the TIMSS 2007 study". Background paper for the *EFA Global Monitoring Report 2010*.

Bhattacharjea, S., W. Wadhwa and R. Banerji (2011). *Inside Primary Schools: A Study of Teaching and Learning in Rural India.* Mumbai: ASER and Pratham Education Initiative.

Burger, R. (2011). "School effectiveness in Zambia: The origins of differences between rural and urban outcomes". *Development Southern Africa.*Vol. 28, No. 22, pp. 157-176.

Carr-Hill, R. (2017). "Improving Population and Poverty Estimates with Citizen Surveys: Evidence from East Africa". *World Development*. Vol. 93, pp. 249-259.

Cueto, S., G. Guerrero, J. Leon, M. Zapata and S. Freire (2014). "The relationship between socioeconomic status at age one, opportunities to learn and achievement in mathematics in fourth grade in Peru". *Oxford Review of Education*. Vol. 40, No. 1, pp. 50-72.

Educational Initiatives (2010). *Student Learning Study: Status of Student Learning across 18 States of India in Urban and Rural Schools*. Ahmedabad Educational Initiatives. Available at http://www.ei-india.com/student-learning-study-an-india-report/

Fehrler, S. and K. Michaelowa (2009). "Education marginalization in sub-Saharan Africa". Background paper for the *EFA Global Monitoring Report 2010*.

Filmer, D. and L. Pritchett (1999). "The effect of household wealth on educational attainment: evidence from 35 countries". *Population and Development Review.* Vol. 25, No. 1, pp. 85-120.

Filmer, D. and K. Scott (2012). "Assessing asset indices". *Demography.* Vol. 49, No. 1, pp. 359-392.

Goldstein, H. (1983). "Measuring changes in educational attainment over time: Problems and possibilities". *Journal of Educational Measurement*. Vol. 20, No. 4, pp. 369- 377.

Jones, S. and Y. Schipper (2012). "Does family background matter for learning in East Africa?" *Africa Education Review*. Vol. 12, No. 1, pp. 7-27.

Lewin, K. and R. Sabates (2012). "Who gets what? Is improved access to basic education pro-poor in Sub-Saharan Africa?" *International Journal of Educational Development.* Vol. 32 (4), pp. 517-528. https://doi.org/10.1016/j.ijedudev.2012.02.013

Moloi, M. Q. and M. Chetty (2010). *The SACMEQ III Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education*. Pretoria: South Africa Ministry of Basic Education.

Muralidharan, K. and Y. Zieleniak (2013). "Measuring learning trajectories in developing countries with longitudinal data and item response theory". Presentation for Young Lives Conference, Oxford. Available at http://www.fsmevents.com/younglives/session04/slides.pdf

Newton, P. (1997). "Examining standards over time". *Research Papers in Education*. Vol. 12, No. 3, pp. 227-248.

Pritchett, L. and A. Beatty (2015). "Slow down, you're going too fast: Matching curricula to student skill levels". *International Journal of Educational Development*. Vol. 40, No. 2015, pp. 276-288.

Rolleston, C., Z. James and E. Aurino (2014). "Exploring the effect of educational opportunity and inequality on learning outcomes in Ethiopia, Peru, India, and Vietnam". Background paper for *Education for All Global Monitoring Report 2013/4*.

Rose, P. (2015). "Three lessons for educational quality in post-2015 goals and targets: Clarity, measurability and equity". *International Journal of Educational Development*. Vol. 40, pp. 289-296.

Rose, P. and B. Alcott (2015). *How Can Education Systems become Equitable by 2030? DFID Think Pieces – Learning and Equity*. Report prepared for UK Department for International Development. Oxford, UK: Health and Education Advice and Resource Team.

Rose, P., R. Sabates, B. Alcott and S. Ilie (2016). "Overcoming inequalities within countries to achieve global convergence in learning". Background paper for *The Learning Generation: Investing in Education for a Changing World*, International Commission on Financing Global Education Opportunity.

Schreiner, M. (2010). "A simple poverty scorecard for Pakistan". *Journal of Asian and African Studies*, Vol. 45, No. 3, pp. 326-349.

Singh, A. (2014). *Emergence and Evolution of Learning Gaps across Countries: Linked Panel Evidence from Ethiopia, India, Peru and Viet Nam*. CSAE Working Paper WPS/2014-28. Retrieved from http://www.csae.ox.ac.uk/materials/papers/csae-wps-2014-28.pdf.

Smits, J. and R. Steendijk (2015). "The international wealth index (IWI)". *Social Indicators Research.*Vol. 122, No. 11, pp. 65-85.

Spaull, N. and J. Kotze (2015). "Starting behind and staying behind in South Africa: The case of insurmountable learning deficits in mathematics". *International Journal of Educational Development.* Vol. 41, pp. 13-24.

Tayyaba, S. (2012). "Rural-urban gaps in academic achievement, schooling conditions, student, and teachers' characteristics in Pakistan". *International Journal of Educational Management.* Vol. 26, No. 1, pp. 6-26.

UNESCO (2012). *Education for All Global Monitoring Report 2012: Youth and skills – Putting Education to Work*. Paris: UNESCO. http://unesdoc.unesco.org/images/0021/002180/218003e.pdf

—— (2014). *Education for All Global Monitoring Report 2013/4: Teaching and Learning: Achieving Quality for All.* Paris: UNESCO. http://unesdoc.unesco.org/images/0022/002256/225660e.pdf

—— (2016a). *Global Education Monitoring Report 2016: Gender Review: Achievements and Challenges*. Paris: UNESCO. http://unesdoc.unesco.org/images/0024/002460/246045e.pdf

—— (2016b). *Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4 – Towards Inclusive and Equitable Quality Education and Lifelong Learning for All*. Paris: UNESCO. http://unesdoc.unesco.org/images/0024/002456/245656e.pdf

UNESCO Institute for Statistics (UIS) (2017). "Availability and Reliability of Education Finance Data in Household Surveys." UIS Information Paper No. 43. Montreal: UIS. http://uis.unesco.org/sites/default/files/documents/availability-reliability-education-finance-data-household-surveys_0.pdf

United Nations (2007). "*Convention on the Rights of Persons with Disabilities"*. General Assembly resolution A/RES/61/106 (13 December 2006). http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/61/106

—— (2017). "Work of the Statistical Commission Pertaining to the 2030 Agenda for Sustainable Development." General Assembly resolution A/RES/71/313 (6 July 2017). https://undocs.org/A/RES/71/313

United Nations Statistics Division (2017). *SDG Indicators*. Regional groupings used in 2017 Report and Statistical Annex. Available at: https://unstats.un.org/sdgs/indicators/regional-groups/

Watkins, K. (2015). "Leaving no one behind: an agenda for equity". *The Lancet*. Vol. 384, No. 9961, pp. 2248-2255.

5. Redistribution of government financing to promote equity in education

BY SONIA ILIE, PAULINE ROSE AND ASMA ZUBAIRI
Research for Equitable Access and Learning (REAL) Centre, University of Cambridge

INTRODUCTION

Tracking progress towards education goals signals the existence of a political will to address inequality gaps but is not on its own sufficient. Another way to assess true commitment is revealed by the way that government resources are allocated. In many countries, children and young people who are the hardest to reach are often the last to benefit from government investment in education.

Reaching children who are disadvantaged by household income, gender, geographic location, ethnic origin or disability is likely to entail a much larger investment than that required to reach children from groups not affected by disadvantage. The cost is likely to be far higher because the interventions needed to mitigate the root causes of inequalities, which are historically, socially and culturally embedded within societies, are more extensive and complex. As such, while equality of funding means disbursing the same amount of money to each student or school, equity of funding means providing additional resources to those who face inherited disadvantages. Only in this way will every child come to enjoy the same opportunity to achieve their potential in education. We have already seen "justice as fairness" described in Chapter 2. Specifically, it looked in more detail at measuring the extent to which the distribution of education funding compensates for some measure of existing disadvantage, such as poverty or gender.

In Chapter 5 we examine ways to identify who benefits from government education spending, we look at formula funding strategies for redistribution of government financing and we describe how national education accounts can aid the monitoring of how funds are spent to promote equitable outcomes in education.

5.1 ASSESSING WHO BENEFITS FROM GOVERNMENT SPENDING ON EDUCATION[72]

Benefit incidence analysis assesses which population groups benefit the most from government spending on education. While this type of analysis has largely focused on groups defined by income, benefit incidence analysis can also be used to assess

72 Section 5.1 draws on Ilie, S. and P. Rose (2017) "Who benefits from public spending on higher education in South Asia and sub-Saharan Africa?" *Compare: A Journal of Comparative and International Education*, pp. 1-18. https://doi.org/10.1080/03057925.2017.1347870. © British Association for International and Comparative Education. Figures 5.1 to 5.3 were previously published in this article. Reprinted by permission of Taylor & Francis Ltd, on behalf of British Association for International and Comparative Education. Sections 5.2 to 5.4 draw on the authors' background paper for the International Commission on Financing Global Education Opportunity (Zubairi and Rose 2016).

differences in government spending on education by location (rural versus urban or by administrative or geographic sub-region), gender or the proportion of resources that are directly allocated to pupils of different ethnic backgrounds. As such, benefit incidence analysis is a useful tool in measuring whether different groups receive their fair share of government resources.

To carry out cross-country benefit incidence analysis, it is important to have data generated from two main sources. First, nationally-representative data, for example from household surveys, such as DHS, are necessary to obtain estimates of enrolment at each level of education (primary, secondary and tertiary). The same nationally-representative data should also contain information on income or wealth to generate an index that is used to estimate the relative wealth of households in each country and rank individuals according to household wealth (Rutstein and Johnson, 2004). The second source of data is the UIS database, which provides estimates of total government expenditure by level of education in purchasing power parity (PPP) in U.S. dollars. These data are then used to analyse the distribution of government education expenditure to different wealth groups.

The potential to carry out benefit incidence analysis is often restricted in many countries due to a lack of needed data. Ideally the analysis would be undertaken using data on learning outcomes (including children in and out of school) and across different equity dimensions (e.g. household wealth, location, gender and disability). The analysis can further assess the distribution of resources to different types of schools attended by children with various backgrounds. In other words, the analysis would show the relative amount of funding received by children from an advantaged background compared with those from a disadvantaged background, adjusted for whether or not they are in school. **Box 5.1** illustrates how to perform a benefit incidence analysis.

It is likely that many countries have more detailed data than those collected by international agencies

Box 5.1 Data needs for benefit incidence analysis

The main types of data required for benefit incidence analysis include:

- Enrolment (or attendance) by level of education - disaggregated by sex, location, wealth, disability and other markers of disadvantage. Most countries should have such information from a national household survey.
- Learning outcomes by level of education.
- Government expenditure on education, by level of education.

Using enrolment data, a formalisation of the analysis for each level of education is as follows. In the case of the analysis presented in this section, this includes three levels of education, namely primary, secondary and tertiary (adapted from Demery, 2000):

$$X_j = \sum_{i=1}^{3} E_{ij} \frac{S_i}{E_i} = \sum_{i=1}^{3} \frac{E_{ij}}{E_i} S_i \qquad (1)$$

where X_j represents the total value of government education expenditure for group j, which is indicated by the level of disaggregation indicated above. E_{ij} refers to total enrolment of group j at education level i. S_i is government expenditure for education level i, and E_i represents total enrolment in education level i.

Equation (1) can also be used to examine government expenditure for each level of education to determine which populations benefit the most and to assess whether the distribution of government resources is less equitable for tertiary education than for primary and secondary education.

to allow more nuanced analyses. Such analyses can take account of spending by students, their parents or other organizations to enable access to education. This would normally include any fees charged at different levels of public education. If it is assumed that the same fees are charged to rich and poor families alike in all public institutions, government expenditure for each of these groups will remain accurate. The analysis can also take account of the extent to which students from different backgrounds pay fees to attend private schools. The analysis should be more systematically organized at the country level, with international agencies supporting data collection in ways that can inform the distribution of education spending.

The analysis described above assumes that all those attending school are enrolled in public institutions and that they either do not pay fees or pay the same level of fees. However, if the analysis does not take into account the out-of-pocket costs of education (e.g. uniforms, textbooks or examination fees) paid by parents, the results may underestimate the potential inequalities in the distribution of resources. Conversely, if wealthier children are more likely to attend privately-funded institutions – and so incur the costs of provision themselves rather than receiving government funding – we would overestimate the amount of government subsidies to wealthy households and thus the degree of inequality in funding allocations, because wealthier children receive in fact little or no government subsidy. As a result, the actual amount spent by the government per child from a rich household will be lower than the estimates presented in this section.

Overall, the patterns in inequality in funding are often not straightforward. For example, wealthier households might invest in better-quality private schooling at lower levels of education, enabling their children to benefit from larger subsidies in public institutions at higher levels. In addition, certain countries use government spending to subsidise attendance at private schools, whereas other countries provide grants to all schools, whether public or private. For comprehensive analyses of the association between enrolment patterns and the distribution of public and private spending on education, data on government expenditure on education, private spending and enrolment by household wealth level at each level of education are needed.

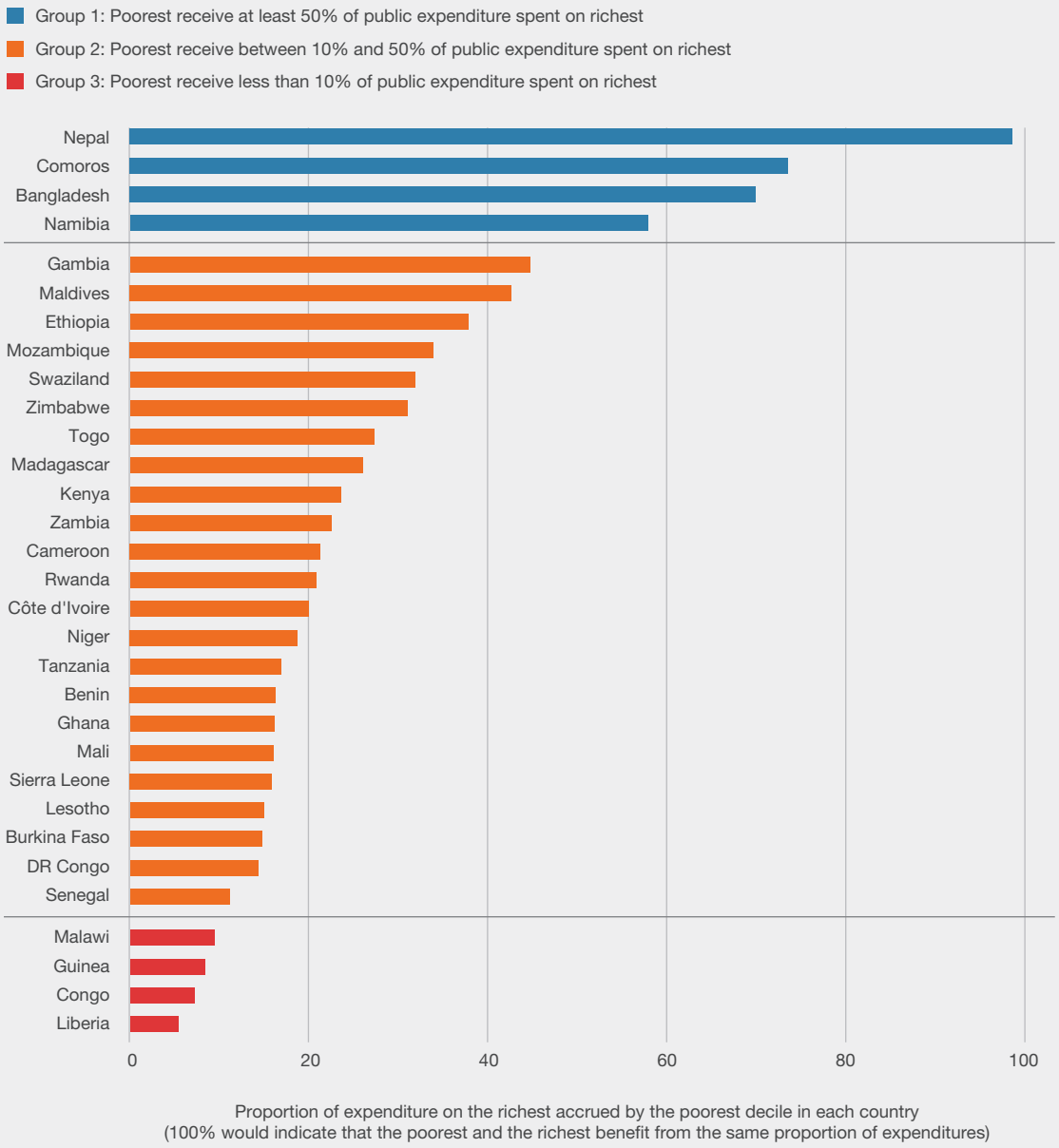
Despite government expenditure on education being a potential source of resource redistribution, benefit incidence analysis shows that there are still large inequalities in the distribution of government spending on education within countries. To illustrate this point, **Figure 5.1** shows vast wealth gaps in government spending within 31 low- and lower-middle-income countries in sub-Saharan Africa and South Asia for which data were available. While government education expenditure disproportionately benefits the richest groups in all 31 countries, the levels of inequity vary substantially. In Bangladesh, Comoros, Namibia and Nepal, for instance, for every $100 spent on children living in the richest 10% of all households, at least $50 are spent on children from the poorest 10%. At the other extreme, in Congo, Guinea, Liberia and Malawi, children from the poorest decile benefit from less than $10 for every $100 spent on children living in the richest 10% of households (Ilie and Rose, 2017).

It is important to highlight that spending $100 on both a rich and poor child would amount to an equal distribution of spending. It would not, however, be equitable. An equitable distribution would require a larger proportion to be allocated to children from more disadvantaged backgrounds who require a greater share of resources so that gaps in enrolment and learning due to inherited disadvantage can be narrowed.

Further insights from benefit incidence analysis can be gained by comparing the distribution of government spending on education by levels of education. Using information from 31 low- and lower-middle-income countries in sub-Saharan Africa and South Asia, Ilie and Rose (2017) find that government expenditure on primary education is pro-

Figure 5.1 In most of 31 selected low- and lower-middle-income countries, the poorest children receive just a fraction of government expenditure on education

Government education spending on the poorest household decile relative to spending on the richest household decile (%), in descending order, latest year available



Note: As far as possible, the data have been matched for the years available for expenditure data from the UIS with the years for the DHS school attendance data, using the most recent year available. The years of data available ranged from 2005-2014. Data on government expenditure for all levels of education were analysed.
Source: Ilie and Rose (2017) based on calculations with UIS and DHS data.

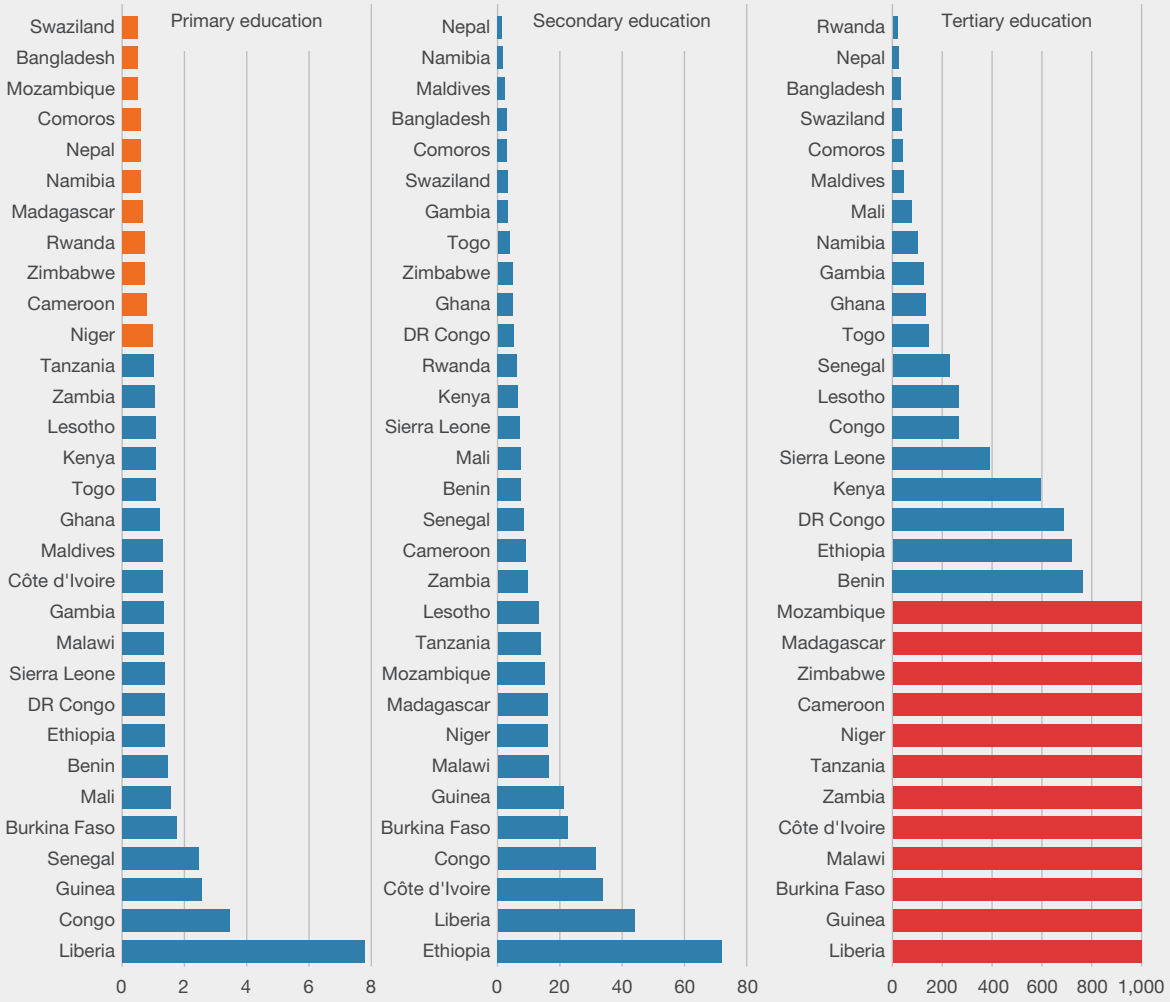
poor in one-third of countries, with children from the poorest 10% of households benefitting from larger shares of government resources than children from the richest 10% of households *(see **Figure 5.2**)*. However, government expenditure on secondary education is pro-rich for all countries, although there is great variation among them. For example, in Nepal, government spending on secondary education for the richest decile is just 1.3 times the amount for the poorest decile, while in Ethiopia, government spending on secondary education for children from the richest 10% of households is 72 times the

Figure 5.2 The poor-rich disparity in beneficiaries of government education expenditure is most extreme in tertiary education

Ratio of government education spending on the richest decile relative to spending on the poorest decile, by level of education, latest year available.

A ratio smaller than 1 indicates that the richest household decile accrues less benefit from government expenditures than the poorest decile.

Countries where the richest:poorest ratio significantly exceeds 1,000.



Note: As far as possible, the data have been matched for the years available for expenditure data from the UIS with the years for the DHS school attendance data, using the most recent year available. The years of data available ranged from 2005-2014.
Source: Ilie and Rose (2017) based on calculations with UIS and DHS data.

amount spent on children from the poorest 10%. At the tertiary level, government spending is extremely skewed in favour of the rich for all countries. Most strikingly, Figure 5.2 shows that in 12 of the 31 countries, government expenditure on tertiary education is 1,000 times more for those from the richest 10% of households than for the poorest 10%. The main reason for this discrepancy is that so few of the poorest gain access to tertiary education.

Overall, benefit incidence analysis for the 31 low- and lower-middle-income countries shows that countries with the largest inequalities in government education expenditure on rich versus poor households at the tertiary level also demonstrate more pro-rich expenditure patterns at the primary and secondary levels. For example, Ghana distributes a larger share of education expenditure to tertiary education and also exhibits higher overall inequalities, and Malawi is both highly unequal in the total distribution of education expenditures and spends a large proportion of this on tertiary education. By contrast, Nepal displays both low inequalities and low expenditure on tertiary education relative to other levels of education (Ilie and Rose, 2017).

A further representation of inequalities in the distribution of government spending on education can be obtained using a concentration curve (defined in Chapter 2, Section 2.3.3 and illustrated in Chapter 3). The concentration curve represents an approach to analysing impartiality, in this case of government education expenditure with respect to household wealth. Specifically, it shows the share of the distribution of government spending on education allocated to each wealth group. Nepal's concentration curve is very close to the 45-degree line, which represents parity across the wealth deciles – i.e. each wealth decile receives around 10% of the share of government spending (see ***Figure 5.3***). In other words, government spending on education in Nepal is nearly equally distributed among the different wealth groups of the population. As discussed in Chapter 2, this does not, however, imply an equitable distribution of resources, which would require more spending on the poorest, resulting in a concentration curve above the 45-degree parity line.

Equal distribution of government spending is not the norm for most low- and lower-middle-income countries, and it is certainly not equitable distribution. In Ghana specifically, and the sub-Saharan African group of countries on average, government spending is unequal: children living in the richest 10% of households receive around the same amount of government spending as children from the poorest 45% of households. Malawi shows an even more dramatic pro-rich distribution, with children in the wealthiest 10% of households receiving the same amount of government spending as children in the poorest 80% of the population. Moreover, children in the richest 10% of households alone received 44% of total public resources for education.

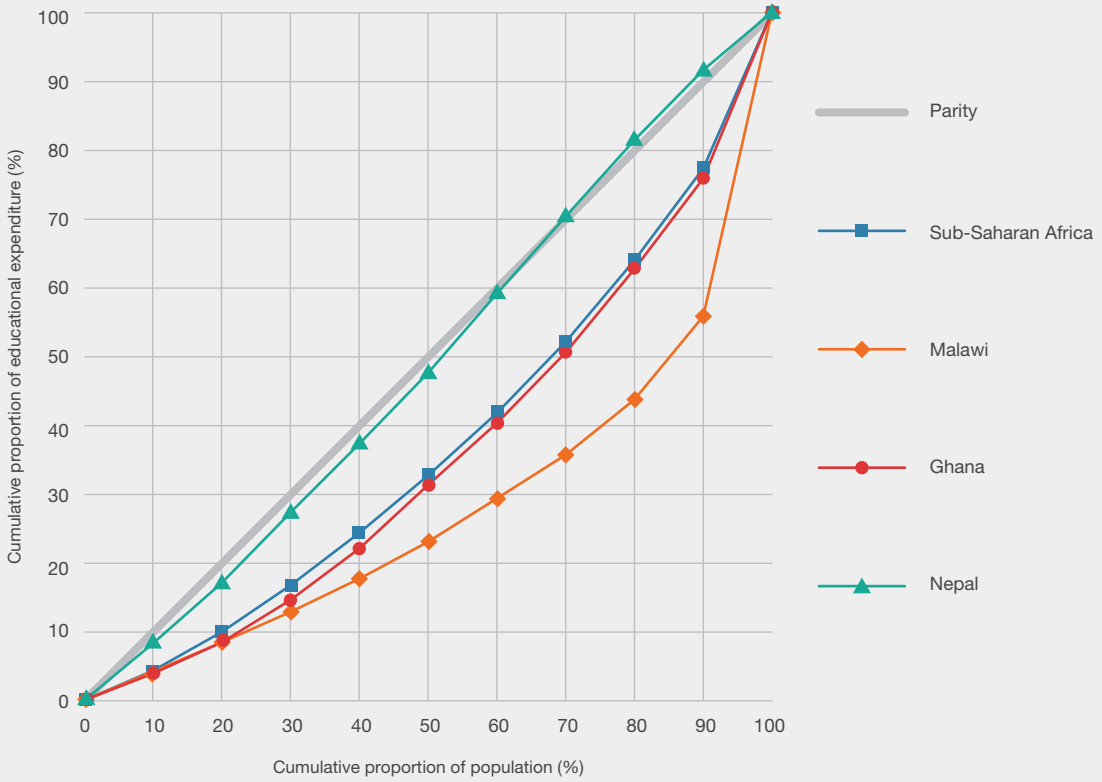
5.2 FORMULA FUNDING STRATEGIES FOR REDISTRIBUTION OF GOVERNMENT EDUCATION RESOURCES

Having ascertained that the distribution of government education resources is often highly inequitable, the question is: how can this be changed? Using a formula that distributes a larger proportion of resources to those most in need is an important policy tool. It directly addresses thematic SDG Indicator 4.5.3, which aims to identify the extent to which "explicit formula-based policies reallocate education resources to disadvantaged populations" (UIS, 2016). The UIS notes that there is currently a lack of clarity in the wording of this indicator, a misalignment between the target and its operationalisation, thus its feasibility is limited.[73] These technical concerns are important but do not undermine the importance of identifying

73 At its second meeting in 2016, the Technical Cooperation Group on the Indicators for SDG 4-Education 2030 identified SDG Indicator 4.5.3 as one of the indicators requiring further methodological development.

Figure 5.3 Pro-rich biases in government expenditure vary greatly across selected countries

Cumulative distribution of total government educational expenditure by wealth decile in Nepal, Ghana, Malawi and sub-Saharan Africa



Note: Data refer to 2008 for Ghana, 2010 for Malawi and 2009 for Nepal. Data on government expenditure for all levels of education were analysed.
Source: Ilie and Rose (2017) based on calculations with UIS and DHS data.

whether and how such formula-based policies are working, as this chapter illustrates.

This section provides an overview of approaches to formula funding as adopted by some countries. As the section highlights, it appears that these formulae often do not sufficiently tackle inequities for a variety of reasons. Improved measurement of who benefits from government spending, as outlined in Section 5.1, would enable a better understanding of the effects of different funding formulae and of whether attempts to improve them have the desired effect.

Since the 1980s many governments around the world have moved towards a system of allocating funding for education through a per-pupil formula (Levacic, 2014). These formulae are often based on the principle of equality of funding, where funding is allocated on the basis of the numbers of children in school but do not factor in the differing needs of providing education to harder-to-reach groups who are more likely to be out of school. By failing to differentiate between the backgrounds of students in different locations, equality of funding approaches can be highly regressive. As highlighted in this handbook, equitable funding approaches are based on the principle of allocating different amounts of money per child based on need. While some countries have introduced formulae to allocate funding in this way, it often makes up only a small proportion of total

government education funding, particularly where teacher salaries are excluded from the formula.

In addition, where they exist, these formulae usually focus on access, failing to take account of the need to redistribute further to overcome wide inequalities in learning between better-resourced and poorly-resourced schools, with children from disadvantaged backgrounds most often finding themselves in the latter. The message from this is that adopting formulae for the distribution of government resources is an important first step but unlikely to be sufficient on its own. The experience from countries that have specifically targeted disadvantage through redistribution policies offers important practical lessons on how their design can be strengthened to improve the learning outcomes of the most disadvantaged children.

5.2.1 Examples of formula funding

A number of middle- and high-income countries have embarked on differentiated funding according to need, with varying degrees of success and different approaches to addressing inequity in education. In the Netherlands, for instance, a school funding formula in place since 1985 is weighted in favour of the number of disadvantaged children in a primary school. The main categories of disadvantaged children are defined as native Dutch children whose parents themselves have little education and disadvantaged immigrant children. Schools with a large number of disadvantaged children, on average, have 58% more teachers per student, together with more support staff (Ladd and Fiske, 2010).

Some of the highest inequalities in the world continue to persist in middle-income countries with large populations, where the benefits of impressive economic growth rates over the last two decades have been very unequal. Some of these countries are now attempting to reverse the trends by distributing education resources through a funding formula that favours disadvantaged groups, whether this is by poverty level, geographic region or school *(see **Table 5.1**).*

Some low- and lower-middle-income countries have also begun to consider ways in which education resources can be redistributive, although their experience is not as advanced as in some of the middle-income countries. Part of this is due to the capacity constraints that these governments face in targeting and implementing programmes effectively.

In Kenya, for instance, the government adopted a new constitution in 2010, which set out how government allocations would meet the requirements relating to equity. It was aimed at reducing disparities between regions and closing the gap in relation to access and quality of basic services, including education. Article 202 of the 2010 constitution sets out that revenue raised nationally should be shared equitably among national and county governments. The equitable share provision under Article 203 stated that counties would receive a minimum of 15% of national revenue. A further 0.5% of revenue would be channelled by the Equalisation Fund, created in 2010 by the constitution to improve services in the most marginalised parts of the country to "bring the quality of those services […] to the level generally enjoyed by the rest of the nation, so far as possible" (Watkins and Alamayehu, 2012). However, capitation grants to primary schools have been dependent on the number of children enrolled, which not only fails to take into account where disadvantaged groups may be in need of more resources but also penalises the poorer arid and semi-arid regions of the country, which are home to 46% of the out-of-school population (Watkins and Alamayehu, 2012).

In Malawi, schools currently receive two grants directly: the "Direct Support to Schools" grant initiated in 2006 and the much larger "School Improvement Grant" introduced as a pilot in 2010. The latter focuses specifically on orphans, vulnerable children and HIV-positive children. This grant is disbursed at school level on the basis of both the total number

Table 5.1 Selected middle-income country interventions to increase equity of education expenditure

	Name of scheme	Description
Brazil	*Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização de Profissionais de Educação* (**FUNDEB)**	In 1996, Brazil introduced the *Fundo para Manutenção e Desenvolvimento do Ensino Fundamental e Valorização do Magistério* (**FUNDEF** - Fund for Primary Education Administration and Development and for the Enhancement of Teacher Status) with the aim of investing a minimum amount per child, with state spending complemented by federal allocations. In 2006, FUNDEF became *Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização de Profissionais de Educação* (**FUNDEB** - Fund for Basic Education Administration and Development and the Enhancement of the Status of Education Professionals). It further differentiated between categories of students to account for certain marginalised groups. In addition, it was extended beyond primary education to include primary and secondary levels and youth and adult education.
India	Special Focus Districts under *Sarva Shiksha Abhiyan* **(SSA)**	*Sarva Shiksha Abhiyan* (SSA – Universalisation of Elementary Education) has been operational since 2000 and is intended to reduce the gaps in education outcomes between regions, income groups, gender and minority communities. In 2006, SSA identified districts to receive additional funding based on variables including out-of-school population, gender disparities and large populations of minority groups. These were Special Focus Districts.
Mexico	Compensatory Education Programme implemented by *Consejo Nacional De Fomento Educativo* **(CONAFE)**	In 1992, the government, through the *Consejo Nacional de Fomento Educativo* (CONAFE - National Council of Education Promotion), implemented a Compensatory Education Programme. The programme aimed to ensure an equitable distribution of resources and education standards across all schools.
Pakistan	National Finance Commission Award	Federal transfers account for 90% of provincial spending on education. Until 2009, funding to provinces was disbursed according to the share of their population. As this failed to take into account the deprivation of provinces and their ability to raise their own resources it was considered a regressive funding formula. The 7th Award in 2009 added three additional criteria to determine provincial allocations. This took into account disadvantage of certain provinces in mobilising resources and poverty levels.
South Africa	National Norms and Standards for School Funding	The programme was introduced in 2006 to reverse the legacy of Apartheid. It was intended to reduce the wide gaps by race in education spending.
Sri Lanka	Education Quality Inputs	The programme was introduced in 2000 to address disadvantage faced by smaller schools in rural and poorer areas which have higher operating costs.
Viet Nam	Primary Education for Disadvantaged Children Project	The project was implemented from 2003 to 2010, focusing on achieving fundamental school quality levels with particular attention to initiatives for highly-vulnerable children. 227 of the poorest districts targeted out of a total of 615 districts. This targeted approximately 70% of disadvantaged children.

Sources: Zubairi and Rose (2016), drawing on Da Cruz et al. (2015); Bruns et al. (2012); Jhingran and Sankar (2009); Motala, Dieltiens and Sayed (2012); Giese et al. (2009); Arunatilake and Jayawardena (2013); Getler et al. (2012); Poisson (2014); Mestry and Ndhlovu (2014); Malik and Rose (2015).

How does the scheme direct resources to most disadvantaged?	What does the scheme cover?	Shortcomings of current funding formula approach
Federal funding supplements state funding to reduce historical disparities in education funding between poorer and richer states, and to counter the highly unequal tax mobilisation capacity among states. This funding formula favours rural and marginalised indigenous groups. The major part of funding is earmarked for teacher salaries in order to compensate poorer states in the North and North-East that historically had a larger share of unqualified teachers.	FUNDEB resources earmark 60% for teacher salaries and 40% for school operations.	
Government disburses additional resources to states/districts with large out-of-school populations, gender disparities, disadvantaged minority groups and children with special needs.	SSA disburses grants to schools under these broad headings: i) maintenance; ii) development; and iii) teaching and learning.	• Grants to schools are earmarked under three headings only which may not be aligned with school needs. • There is no mechanism to compensate schools for increased enrolment as grants are not based on per-child norms. • Large schools, with higher operating costs, are penalised, meaning total grant per child allocation is inequitable when comparing between types of school.
Funds skewed in favour of the poorest districts within the 23 poorest states. Spending is on infrastructure improvement, school equipment, textbooks and teaching and learning materials, pedagogical training, and monetary incentives for teachers. In addition, small amounts of funds are disbursed to parents who have the discretion to invest them in what they deem appropriate for the needs of the school.	Non-salary expenditure including monetary incentives for teachers and support for school management committees.	• Excludes teacher salaries and capital spending.
Federal transfers to provinces make up 90% of total education resources for provinces. The distribution of these federal resources to the provinces is weighted on four criteria: 82% is on population size, 10% on "poverty backwardness", 5% on provincial revenue collection, and 3% on low population density. The change in formula is intended to favour smaller, sparsely-populated and less-developed provinces.	All education expenditure.	• Weighting given to poverty is just 10%. • Formula-based budgeting within the four provinces is currently not operational despite intra-province inequalities.
Fee-free schooling in geographical areas with the poorest three income quintiles. Each Provincial Education Department directs 60% of non-personnel and non-capital recurrent expenditure towards schools in the bottom two income quintiles.	Non-salary expenditure.	• Excludes teacher salaries and capital spending. • Funding is earmarked and so may not align with school needs. • A school might be situated in a wealthy catchment area and yet the learners might come from low-income households. Similarly, learners from wealthy households may be attending no-fee schools without having to make any contribution.
Funds are skewed towards smaller schools which are largely situated in rural and poorer areas. Resources are earmarked for purchase of consumables and repair and maintenance of schools.	60% for consumables 40% for repair and maintenance	• Excludes teacher salaries. • Funding is earmarked and so may not align with school needs.
The project targeted disadvantaged children, i.e. those not enrolled in school, attending sub-standard schools, being disabled or belonging to other vulnerable groups. Targeting was at the district level.		• Targeting based on district and not school. Within a district all schools are eligible to receive project support, meaning some better-off schools could be included.

of children and the number of vulnerable children enrolled. At the national level, a per child amount of US$12.90 was set for orphans and vulnerable children, while US$20 was the flat rate for HIV-positive learners (Nampota and Chiwaula, 2014).

5.2.2 Formula funding needs to take account of both access and quality

While increasing attention has been given to the distribution of education resources according to enrolment, there has been less focus on how this distribution affects the quality of education in schools that the majority of poor children attend. In South Africa, for instance, 60% of the poorest learners are three grade levels behind learners in the richest quintile by Grade 3. This gap increases to four grade levels by Grade 9 (Spaull and Kotze, 2015). South African schools with an overwhelmingly large population of learners of African descent still underperform schools with learners who are predominantly of Afrikaans or Asian descent. In spite of additional resources being channelled more equitably according to the National Norms and Standards for School Funding policy, the increase in the number of candidates passing examinations successfully has been minimal (Mestry, 2014).

As such, in the context of extremely wide inequalities that are reinforced within schools, formula funding alone is unlikely to be enough to close inequality gaps. Despite some countries with large inequalities, such as Brazil, India and South Africa, institutionalising a redistribution of government education resources towards the most marginalised and disadvantaged parts of the country, and in some cases even managing to reduce the gap in learning disparities, there remains a large divide in both the funding for and learning outcomes in schools and regions deemed as poor compared to their rich counterparts. For example, in Brazil, schools with a large number of disadvantaged children are also characterised as having fewer qualified teachers, worse infrastructure and fewer contact hours for teaching (Simoes and Sabates, 2014).

An additional challenge is that, even where government resources have been skewed in favour of poorer schools, wealthier schools continue to be able to supplement government resources with other sources of income. In Sri Lanka, for instance, despite non-salary recurrent expenditure being skewed in favour of schools from the bottom three income quintiles, schools in the top two quintiles – where state funding has been reduced – continue to be able to acquire physical and human resources paid for through school fees, thereby perpetuating the inequities in funding (Mestry and Ndhlovu, 2014).

5.2.3 Formula funding needs to take account of regional inequalities within decentralised systems

In large decentralised systems, where funding for education is determined by both local governments and the central government, inequalities may be perpetuated given that poorer states will rarely be able to contribute as much as wealthier states. In India, for instance, government expenditure per child in elementary school in 2011-2012 was approximately US$126 and US$256 for the wealthier states of Kerala and Himachal Pradesh, but only US$79 and US$54 respectively for West Bengal and Bihar (Dongre et al., 2014).

Similarly, in Pakistan, there is wide variation in the capacity of different provinces to raise revenue for education through local taxes. The wealthier province of Punjab raises 11% of its total revenue through its own taxes. The equivalent for the poorer province of Baluchistan is just 3%. In spite of the changes brought in to ensure more equitable resource distribution among the provinces in Pakistan in 2009, the amounts that provinces disburse to the districts are still beset by large inequities. In all four provinces of Pakistan, education budgets are the lowest in the

districts that have the highest proportions of out-of-school children (Malik and Rose, 2015).

5.2.4 Redistributive formulae need to include teacher salaries

In many countries where a funding formula is in place, salaries for teachers – which make up by far the largest part of education expenditure – are not included in funding formulae, limiting the effectiveness of redistribution. In Sri Lanka, for instance, the Education Quality Inputs programme makes up 2% of the total education recurrent budget as it does not take account of teacher salaries. Similarly, in South Africa, grants meant to disburse more funds to disadvantaged areas are limited to non-salary recurrent expenditure. Non-salary expenditure makes up only 8% to 10% of school budgets. The distribution of qualified teachers continues to be skewed towards wealthier schools while neglecting schools with a higher pupil-teacher ratio in townships and rural areas (Mestry and Ndhlovu, 2014). Brazil provides a counter example to this, with 60% of FUNDEB funds being earmarked for teacher salaries *(see Table 5.1).* FUNDEB made it mandatory for all teachers to acquire minimum qualifications, meaning that by 2002 almost all teachers in the northern disadvantaged areas of the country had acquired the minimum teaching qualifications required.

5.2.5 Schools need autonomy over the spending of resources, with guidance for using it in ways that address education quality for disadvantaged groups

How resources are used at the school level will determine whether redistribution makes a difference for closing inequality gaps. In India, SSA funding *(see Table 5.1)* prioritises investment for infrastructure over improvement in education quality. Interventions earmarked for "quality" within India's SSA grant are made up of two components: an innovation grant and the Learning Enhancement Programme. However, as a share of elementary school spending, the quality component has been negligible, amounting to less than 1% in financial year 2014-2015 (Accountability Initiative, 2015).

5.2.6 Even where quality issues are addressed at school level, resources rarely reach disadvantaged groups

When robust accountability mechanisms are in place, increasing school autonomy over how resources are spent can make a difference in ensuring that they are used to close inequality gaps. In reality, however, rigid earmarking of how funds can be spent limits the potential effectiveness of decentralised decisionmaking in some countries. In India, despite various commitments made to increase the autonomy of schools and school management committees over spending decisions, this has not happened in practice (Accountability Initiative, 2013). In 2013-2014, for instance, one state expressed an interest in restructuring its in-service teacher training model, while another state asked for a top-grant to improve the quality of textbooks. In neither case were the proposals accepted. Schools and school management committees had no power over how the bulk of total SSA resources were spent (Aiyar et al., 2013). Similarly, in Sri Lanka, education quality inputs are chosen according to guidelines administered by the Ministry of Education. These do not always correspond to what the schools themselves would choose to spend the money on, and the rigid definition of budget lines makes it difficult to change this in order to better accommodate the school's needs (UNESCO, 2011).

The degree of autonomy over spending also creates inequities between the disadvantaged and advantaged. In Sri Lanka and South Africa, schools that receive the bulk of their funding from Education Quality Inputs and National Norms and Standards for School Funding *(see Table 5.1)* – which are the schools that the most disadvantaged students are likely to attend – have less autonomy over spending than schools that can mobilise additional income from

school fees and other sources of income. In South Africa, richer schools are able to hire teachers using other resources, in addition to funds they receive from the government. Given that South Africa's National Norms and Standards for School Funding apply only to non-recurrent expenditure, this means that poorer schools, whose income is completely dependent on government funds, are unable to spend them on teachers even where this is the greatest need. The inequities over school autonomy and how school grants are spent also exist in Honduras between less- and more-developed regions. In less-developed regions, it was found that parents had no power over recruitment of teachers, whereas in more-developed areas parents had more autonomy over teacher recruitment.

Greater school autonomy can nonetheless improve learning, as the example of Mexico illustrates. Schools in Mexico were able to double the grants distributed to parent associations, and this was associated with significant improvements in test scores in Spanish and mathematics (Getler et al., 2012).

Based on these country experiences, a first recommendation would be to ensure that grants provided to schools are not rigidly earmarked. Doing so often means that the true needs of schools receiving these funds go unserved. Second, any grants offered should be accompanied by increased support to improve the capacity of schools and school management committees to ensure that they are able to use the grants in ways that will improve learning outcomes and thus close inequality gaps.

5.2.7 The effects of redistribution on narrowing learning gaps take time

Seeing the effects of the redistribution of resources towards poorly-endowed schools in order to close learning gaps is likely to take time. In Brazil, learning gaps between public schools with less than 20% disadvantaged pupils and public schools with 80% to 100% disadvantaged pupils decline as a function of the length of time during which support is provided, as well as the amount of resources given (Simoes and Sabates, 2014). This suggests that although attainment gaps continue for a short period of time (over the first two years), a commitment by the government to providing substantial additional resources to marginalised students has the potential to narrow learning gaps after the initial one to two years of financial support.

For this to happen, the flow of resources needs to be regular and reliable. However, a mixture of poorly-timed disbursements and insufficient information on how much schools can expect to receive has impeded progress towards equity, in which case poorer states and schools remain the most affected.

In South Africa, schools in the poor region of the Eastern Cape were receiving smaller allocations than what they were entitled to compared to other regions (Sayed and Motala, 2012). In Sri Lanka, school census data from 2011 indicate that less than one-third of schools had received Education Quality Inputs funds halfway through the school year; in the poorer North Western province just 6% of schools had received these funds halfway through the school year. In both South Africa and Sri Lanka, this late delivery has proved most detrimental for the poorest, given the greater reliance of schools in poorer districts on such funding (Aruntilake and Jayawardena, 2013). Delays have often meant that when money finally does arrive resources cannot be fully utilised, or when they are, the pressure of having to spend them within the financial year is often not aligned with school needs and planning. In India, for example, just 62% of SSA funds were spent in 2011-2012 (Accountability Initiative, 2013). This disguises the much lower utilisation rates among poorer districts and schools. These delays in receiving funds thus frequently lead to an inability to spend them on activities that reflect school needs: instead, they are spent on such items as whitewashing school walls, which allows schools to use these resources before the financial year ends and no further spending is allowed (Aiyar at al.,

2013). Despite these difficulties, evidence from Sri Lanka illustrates how schools with qualified teachers and head teachers are often more adept at using and aligning funds with school needs, even when there is a delay in receiving them (Arunatilake and Jayawardena, 2013).

Ensuring greater efficiency and effectiveness of funding will be crucial if equalisation of opportunity is to be achieved. While Sri Lanka has introduced a policy of allowing schools to carry over unspent funds into the next financial year, this fails to adequately address the long-term issue of reduced capacity of poorer schools to spend funds effectively. In Mexico, to counter-balance the problem of delayed disbursements, parents have received skills training in the management and spending of funds (Getler et al., 2012). One recommendation would be to train personnel, including teachers, in financial management. It would also be necessary to improve reliability and information on when funds will arrive. For Kenya, a recent study suggested that the Ministries of Finance and Education establish a disbursement schedule to make funding more predictable; 50% of resources are meant to be disbursed in the first school term, 30% in the second term and 20% in the third term (Njihia and Nderitu, 2014).

In decentralised settings, funds often come from multiple sources, which means that additional pressure is placed on poor schools with limited capacity to spend effectively. According to a study of Indonesia (World Bank, 2013), schools in the country received funds from eight different sources and four different budgets, making the education funding system incredibly complex and equalisation of resources more difficult. In India, a system with multiple agencies that disburse funds to districts and schools has meant that equalisation of resources between districts has been extremely limited (Accountability Initiative, 2013). In Uganda, the large number of conditional grants has resulted in a highly-fragmented system, which makes it difficult to assess whether funding to schools is equitable. In 1997-1998, the number of sector conditional grants was 10, before increasing to 46 in 2014-2015. There was a proposal to reduce the number of conditional grants to 13 in 2015-2016 (Government of Uganda, 2015).

In summary, redistributive measures to equalise education opportunity for the most disadvantaged are important and need to be bolstered by a range of other measures to ensure that funds are not just equitable, but that their effectiveness and efficiency are also increased. First, on the expenditure side, investments must be better linked to interventions that would improve learning outcomes. This would mean less rigid earmarking of education expenditure so as to recognise school autonomy and better address school needs at local levels. Second, given that the largest expenditure item of many country budgets is teacher salaries, it is vital that any funding formula be accompanied by effective deployment of qualified teachers to those schools and regions that are most disadvantaged. Third, it is necessary to improve the disbursement of funds so that schools receive them in a timely and predictable manner. Governments should avoid complex mechanisms for disbursing funds to schools and strive to ensure transparency about what schools might expect to receive. Fourth, schools in more disadvantaged communities need capacity development to support them in the management and use of the funds.

5.3 HOUSEHOLD SPENDING ON EDUCATION: IMPLICATIONS FOR EQUITY

The low levels of current tax collection efforts in many of the poorest countries has translated into inadequate resources for public services, including education. This in turn has led to poor learning outcomes for many of the most disadvantaged groups. Even when governments remain committed to education within their national budgets, they often fall short if the revenue collected from taxes is inadequate to meet education needs (Zubairi and Rose, 2016). As a result, even where free and

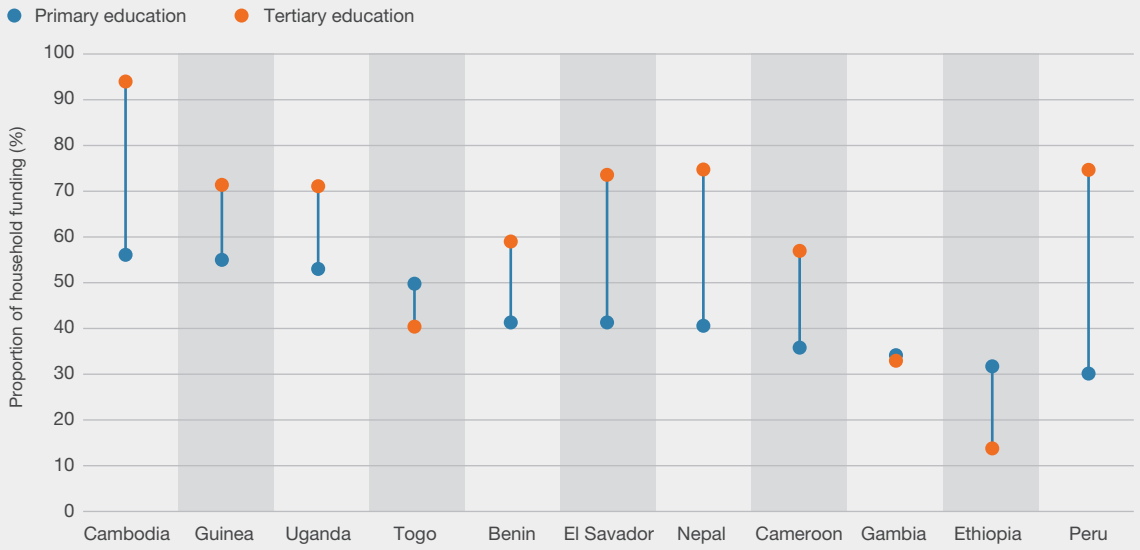
compulsory education has been institutionalised by the constitution, households are often required to make additional contributions, in which case education is not really free. Of 135 countries where fee-free public primary schooling has been guaranteed, 110 continue to charge households a fee of some sort (Transparency International, 2013). Further analysis of data on how much households spend on education is important for assessing the extent to which the distribution of education funding is equitable. Unfortunately, reliable and up-to-date data on household financing of education are currently difficult to obtain.

In many low-income countries, household contributions for sending children to school continue to be significant, especially given the insufficient funding that governments are allocating to the education sector. UIS data reveal the extent to which households are filling the gap by contributing a large share of their resources to supplement government spending on education. In several countries, households contribute 30% or more of combined household and government funding for primary education. In Ethiopia, Gambia and Togo, households are contributing a larger share of spending on primary education than on tertiary education *(see **Figure 5.4**)*. Household expenditure on education as a share of total expenditure (by governments, donors and households) on all levels of education ranged from 14% in Indonesia to 37% in Bangladesh (UNESCO, 2014).

Household contributions to education hit the poorest the hardest, especially in areas where there is low investment by governments. A survey of household spending in 12 African countries found that expenditure on learning materials as a share of total household spending on education was higher among poorer households. On average, such spending consumed 56% of household expenditure on education for the poorest households; yet for the richest households the equivalent was 27%, with a larger share of expenditure on education being spent

Figure 5.4 Households fund a large share of education in some countries

Household contributions as a percentage of combined household and government funding of education per student, in countries where this proportion is 30% or higher (2010-2015 average)



Source: UIS calculations based on UIS data on initial (excluding transfers received) household and government funding per student.

on school fees to send children to private schools (UNESCO Pôle de Dakar, 2012).

5.4 NATIONAL EDUCATION ACCOUNTS FOR TRACKING PROGRESS TOWARDS EQUITABLE FINANCING

One of the routine challenges faced when tracking resources for education in their totality is the lack of holistic, reliable and timely financing data. Many countries do not know how much is being spent on education, on what and by whom. And yet, effective policymaking requires that decisionmakers have access to complete information, including the share of national wealth spent on education; the financial burden of education spending falling on households; the extent to which the education system is supported by external financing sources, such as aid donors or non-governmental providers; and the overall share of education spending on the sub-sectors of education that are accessed by the most disadvantaged.

National Education Accounts (NEA) offer a tool that can help remedy this lack of data and potentially influence finance-related decisions, particularly for vulnerable and marginalised groups. NEAs are a comprehensive, systematic and comparable framework for collecting, processing and analysing data on education expenditure along five dimensions (UIS, IIEP and Pôle de Dakar, 2016):

1) **Financing units**: government (central, state, local), private sector (households, corporations, non-profit organizations), and international sources (bilateral and multilateral donors, foreign, philanthropic organizations).
2) **Producing units**: public and private education institutions, administrative offices.
3) **Level of education**: pre-primary, primary, lower secondary, upper secondary, post-secondary non-tertiary and tertiary education.
4) **Activities**: teaching activities, ancillary services, and general administration and organization of the system.
5) **Economic transactions**: teaching and non-teaching staff compensation, teaching materials, other goods and services, capital expenditure and ancillary services.

In 2013-2016, the UIS, UNESCO's International Institute for Educational Planning (IIEP) and the IIEP Pôle de Dakar implemented an NEA pilot project in eight countries, funded by the GPE.[74] A methodology was developed and published (IIEP, UIS and Pôle de Dakar, 2016) as part of the project, in addition to data collection and publication at the national level. The results were striking, both in terms of collecting data from previously ignored sources and in terms of equity. For example, the NEA project found that:

- Households are major funders of education. They fund about 25% of education expenditure in Viet Nam, around 30% in Côte d'Ivoire, 50% in Nepal and more than 50% in Uganda. In Nepal and Uganda, pre-primary education is almost entirely funded by households, which may have an impact on equity if poorer households are unable to pay.
- Teaching materials are mostly funded by households: 79% of their costs is funded by households in Nepal at the primary level, 66% in Côte d'Ivoire and 50% in Uganda. Again, this has some equity implications as wealthier households have the advantage of being able to invest more than poorer households in these essential inputs for improving learning.
- NEA data can form a strong basis for equity-focused analysis, such as benefit incidence analysis. The national report for Guinea revealed that 10% of the most educated individuals benefit from 39% of government resources allocated to education.

74 Côte d'Ivoire, Guinea, Lao PDR, Nepal, Senegal, Uganda, Viet Nam and Zimbabwe.

Experience from past pilot NEA exercises shows that NEAs can influence policy. A USAID-funded State Education Accounts project in Nigeria indicated a strong bias in Kano and Zamfara states, by both public and private providers, to fund schools in urban areas. As a consequence, state planners reassigned teachers from urban to rural areas. In Zamfara state more funds were channelled towards girls' schools. In addition, more funding was allocated to textbooks and maintenance wherever a shortage of funds was identified (Chawla and Forbes, 2010).

National Health Accounts (NHAs), which have been used for a longer time and much more widely than NEAs, also provide useful lessons. They highlight ways in which publicly-available information has led to widespread support for better health financing decisions, specifically in relation to out-of-pocket expenditures in health, which have adversely affected the most disadvantaged. In Burkina Faso, NHA data revealed huge geographic inequities in health spending, with poorer regions receiving far less than more affluent areas. Boucle du Mouhoun and Nord – two of the poorest regions in Burkina Faso – received a combined 11% of government healthcare spending, in spite of poverty incidences in these regions of 60% and 69%, respectively. By contrast, the wealthier Centre region received 29% of government spending on health, in spite of having an incidence of poverty of only 22%. As a result of these findings, the government and development agencies have now allocated more resources to poorer regions (Zida et al., 2010).

Similarly in India, NHA data revealed that household spending on health accounted for 78% of total health expenditure, thereby making a case for increased government health spending. The large burden borne by households, as revealed in the first round of the NHAs, prompted the government to establish the National Rural Health Mission (2005-2012). Objectives included increasing government expenditure on health and reducing the regional imbalances existing in relation to health infrastructure. It also led to the creation of the *Rashtriya Swasthya Bima Yojna* – a government-funded health insurance scheme targeting the poor (Maeda et al., 2012).

As they plan ahead, education sectors can learn much from the experiences of countries using NEAs. Work on institutionalising and better linking these accounts to high-level policy decisions on education expenditure needs to be accelerated if the Education 2030 goals are to be achieved.

5.5 CONCLUSION

From an equity perspective, there will be little progress towards SDG 4 – inclusive and equitable quality education and lifelong learning opportunities for all – unless resources can be redistributed equitably within education systems according to the principle of "progressive universalism" (International Commission on Financing Global Education Opportunity, 2016). This means that the most disadvantaged should receive the largest share of government resources and pay the least out of their own pockets. NEAs are an important way of tracking progress in this regard and funding formulae are a key mechanism being used across a number of countries for redistribution of education resources. The adoption of redistributive approaches to financing is still nascent in many education systems around the world. Nonetheless, there are already many valuable lessons to be learned. It is vital that all countries pay attention to ensure that no one will be left behind in education by 2030.

REFERENCES

Accountability Initiative (2013). *Do Schools Get Their Money? PAISA 2012*. New Delhi: Accountability Initiative. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2399751

—— (2015). *Do Schools Get Their Money? PAISA 2014.* New Delhi: Accountability Initiative. http://accountabilityindia.in/paisa/study/download/170

Aiyar, Y., A. Dongre, A. Kapur, A. Mukherjee and T. Raghunandan (2013). *Rules vs Responsiveness: Towards Building an Outcome Focussed Approach to Governing Elementary Education Finances in India.* New Delhi: Accountability Initiative. http://accountabilityindia.in/sites/default/files/rules_vs_responsiveness_7.5.15_revised2_0.pdf

Arunatilake, N. and P. Jayawardena (2013). "School Funding Formulas in Sri Lanka". Background paper prepared for the *EFA Global Monitoring Report 2013/14*. http://unesdoc.unesco.org/images/0022/002259/225944E.pdf

Bruns, B., D. Evans and J. Luque (2012). *Achieving World-Class Education in Brazil*. Washington, DC: World Bank. https://openknowledge.worldbank.org/bitstream/handle/10986/2383/656590REPLACEM0hieving0World0Class0.pdf

Chawla, D. and P. Forbes (2010). *Promoting Accountability and Enhancing Efficiency: Using National Education Accounts to Track Expenditure Flows.* Washington, DC: Creative Associates International. http://files.eric.ed.gov/fulltext/ED514809.pdf

Da Cruz, H., A. Luque and A. Teixeira (2015). "Challenges in the Funding Model of State Universities of Sao Paulo". *International Journal of Economics, Finance and Management Sciences*. Vol. 3, No. 3, pp. 148-155. http://article.sciencepublishinggroup.com/pdf/10.11648.j.ijefm.20150303.11.pdf

Demery, L. (2000). *Benefit Incidence: A Practitioner's Guide*. Washington, DC: World Bank.

Dongre, A., A. Kapur and V. Tewary (2014). *How Much Does India Spend per Student on Elementary Education?* New Delhi: Accountability Initiative. http://accountabilityindia.in/sites/default/files/working_paper_series1.pdf

Getler, P., H. Patrinos and M. Rubio-Codina (2012). "Empowering parents to improve education: Evidence from rural Mexico". *Journal of Development Economics*. Vol. 99, No. 1, pp. 68-79. https://openknowledge.worldbank.org/handle/10986/6686

Giese, S., H. Zide, R. Koch and K. Hall (2009). *A Study on the Implementation and Impact of the No-Fee and Exemption Policies*. Cape Town: Alliance for Children's Entitlement to Social Security.

Government of Uganda (2015). "Overview of Reforms: Consolidating intergovernmental transfers to local government" (unpublished document). Kampala: Ministry of Finance, Planning and Economic Development.

Ilie, S. and P. Rose (2017). "Who benefits from public spending on higher education in South Asia and sub-Saharan Africa?" *Compare: A Journal of Comparative and International Education,* pp. 1-18. https://doi.org/10.1080/03057925.2017.1347870

International Commission on Financing Global Education Opportunity (2016). *The Learning Generation. Investing in Learning for a Changing World*. International Commission on Financing Global Education Opportunity. http://report.educationcommission.org/wp-content/uploads/2016/09/Learning_Generation_Full_Report.pdf

International Institute for Educational Planning (IIEP), UNESCO Institute for Statistics (UIS) and IIEP Pôle de Dakar (2016). *Methodology of National Education Accounts*. Paris: IIEP. http://uis.unesco.org/sites/default/files/documents/methodology-of-national-education-accounts-2016-en.pdf

Jhingran, D. and D. Sankar (2009). *Addressing Educational Disparity: Using District Level Education Development Indices for Equitable Resource Allocations in India*. Washington, DC: World Bank. https://doi.org/10.1596/1813-9450-4955

Ladd, H. and E. Fiske (2010). *Weighted Student Funding in the Netherlands: A Model for the U.S.?* Durham, NC: Duke University, Sanford School of Public Policy**.** http://www.iga.ucdavis.edu/Research/EJS/seminars/Ladd%20paper.pdf

Levacic, R. (2014). "Using quantitative data in World Bank per-student funding reform projects: data, designs and dilemmas in transition countries". *Comparative Education.* Vol. 50, No. 3, pp. 340-356. https://doi.org/10.1080/03050068.2014.921373

Maeda, A., M. Harrit, S. Mabuchi, B. Siadat and S. Nagpal (2012). *Creating Evidence for Better Health Financing Decisions: A Strategic Guide for the Institutionalization of National Health Accounts.* Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/13141

Malik, R. and P. Rose (2015). *Financing Education in Pakistan: Opportunities for Action, Country Case Study for the Oslo Summit on Education.* IDEAS Pakistan and University of Cambridge. http://reliefweb.int/sites/reliefweb.int/files/resources/pakista.pdf

Mestry, R. (2014). “A critical analysis of the National Norms and Standards for Social Funding policy: Implications for social justice and equality in South Africa”. *Educational Management Administration & Leadership*. Vol. 42, pp. 851-867. http://journals.sagepub.com/doi/pdf/10.1177/1741143214537227

Mestry, R. and R. Ndhlovu (2014). “The implications of the National Norms and Standards for School Funding policy on equity in South African public schools”. *South African Journal of Education*. Vol. 34, No. 3. http://www.scielo.org.za/pdf/saje/v34n3/02.pdf

Motala, S., V. Dieltiens and Y. Sayed (2012). *Finding Place and Keeping Pace: Exploring Meaningful and Equitable Learning in Southern African Schools*. Cape Town: HSRC Press. https://www.hsrcpress.ac.za/product.php?productid=2307

Nampota, D. and L. Chiwaula (2014). *The Use and Usefulness of School Grants: Lessons from Malawi*. Paris: UNESCO-IIEP. http://unesdoc.unesco.org/images/0022/002282/228286E.pdf

Njihia, M. and J. Nderitu (2014). *The Use and Usefulness of School Grants: Lessons from Kenya.* Paris: UNESCO-IIEP. http://unesdoc.unesco.org/images/0022/002282/228289E.pdf

Poisson, M. (2014). *Achieving transparency in pro-poor education incentives*. Paris: UNESCO-IIEP. http://unesdoc.unesco.org/images/0022/002269/226982e.pdf

Rutstein, S. O. and K. Johnson (2004). “The DHS wealth index”. *DHS comparative reports no. 6.* Calverton: ORC Macro.

Sayed, Y. and S. Motala (2012). “Equity and ‘No Fee’ Schools in South Africa: Challenges and Prospects”. *Social Policy and Administration.* Vol. 46, No. 6, pp. 672-687. http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9515.2012.00862.x/pdf

Simoes, A. A. and R. Sabates (2014). “The contribution of Bolsa Família to the educational achievement of economically disadvantaged children in Brazil”. *International Journal of Educational Development*. Vol. 39, pp. 141-156. https://doi.org/10.1016/j.ijedudev.2014.07.012

Transparency International (2013). *Global Corruption Report: Education*. Berlin: Transparency International. http://www.transparency.org/gcr_education

UNESCO (2011). *Can School Grants Lead to School Improvement? An Overview of Experiences of Five Countries*. Paris: UNESCO-IIEP. http://www.iiep.unesco.org/sites/default/files/school_grants_synthesis_working-paper.pdf

—— (2014). *Education for All Global Monitoring Report 2013/4: Teaching and Learning: Achieving Quality for All.* Paris: UNESCO. http://unesdoc.unesco.org/images/0022/002256/225660e.pdf

UNESCO Institute for Statistics (UIS) (2016). “Improving the international monitoring framework to achieve equity (SDG 4.5): Indicator 4.5.3”. UIS Information Paper No. 32. Montreal: UIS. http://uis.unesco.org/sites/default/files/documents/improving-the-international-monitoring-framework-to-achieve-equity-sdg-4.5-2016-en.pdf

UNESCO Institute for Statistics (UIS), International Institute for Educational Planning (IIEP) and IIEP Pôle de Dakar (2016). *Who Pays for What: The Real Costs Revealed through National Education Accounts.* Montreal: UIS. http://uis.unesco.org/sites/default/files/documents/who-pays-for-what-in-education-national-revealed-through-accounts-2016-en_0.pdf

UNESCO Pôle de Dakar (2012). *Household Education Spending: An Analytical and Comparative Perspective for 15 African Countries*. Dakar: UNESCO-BREDA. http://unesdoc.unesco.org/images/0021/002167/216719e.pdf

World Bank (2013). *Spending More or Spending Better: Improving Education Financing in Indonesia*. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/13207

Zida, A., P. Berone and L. Lorenzetti (2010). *Using National Health Accounts to Inform Policy Change in Burkina Faso.* Health Systems 20/20, Abt Associates. http://pdf.usaid.gov/pdf_docs/pnaeb930.pdf

Zubairi, A. and P. Rose (2016). “Raising domestic resources for equitable education”. Background paper for *The Learning Generation. Investing in Learning for a Changing World*, International Commission on Financing Global Education Opportunity. http://www.educationequity2030.org/resources-2/2017/3/23/raising-domestic-resources-for-equitable-education

6. Concluding remarks

BY PAULINE ROSE, RICARDO SABATES AND BEN ALCOTT
Research for Equitable Access and Learning (REAL) Centre, University of Cambridge

Our world continues to be marked by stark social, economic and structural inequalities. Significant progress towards the SDGs and the Education 2030 Framework for Action is unlikely to take place unless there is strong commitment by UN Member States, global actors and key stakeholders involved in the education sector to promote and deliver equitable, quality education. This handbook was inspired by the need to ensure that equity in education is at the heart of global, national and local agendas aiming to promote access and learning for all children, young people and adults. It also makes the case for greater cooperation and support across governments, donors and civil society for ensuring data are disaggregated by different dimensions of disadvantage to not just track but accelerate progress towards SDG 4.

The handbook has addressed several gaps in the work on equity in education. First was the lack of clarity around the concept of equity in education. Chapter 2 provided a conceptual framework, a set of guiding principles and, importantly, a proposed categorisation that might be considered when applying an equity lens to education. Of particular significance are: the principle of equal opportunities, stating that everyone should have the same opportunity to thrive regardless of the circumstances in which they are born; justice as fairness, which suggests that inequalities are only fair if they can be justified to the least advantaged members of society; and the capabilities approach, which suggests that the set of options a person has open to them is more important than the outcomes they actually achieve.

The handbook has further translated these philosophical arguments into five concepts for measuring equity in education:

a. Meritocracy: educational opportunities are distributed on the basis of merit.
b. Minimum standards: educational opportunities must be at least the same for everyone below a certain threshold.
c. Equality of condition: educational opportunities must be the same for everyone in the population, regardless of their different circumstances.
d. Impartiality: educational opportunities should be distributed equally by gender, ethnicity, religion, language, location, wealth, disability, and other characteristics.
e. Redistribution: mechanism for compensation of initial disadvantage.

A second gap relates to measuring educational equity itself and the operationalisation of educational equity by the many different indicators available. Some of these have emerged as a means to address economic forms of inequalities. The handbook has presented a number of desirable properties that equity measures should meet (Chapter 2) and a detailed explanation of the different equity metrics available from the list of educational indicators (Chapter 3).

A third gap concerns the limitations of the data being collected in education and how much of it excludes information that could be helpful in identifying marginalised groups for a more precise analysis of equity in education. The handbook has provided a sequence of steps to be followed with currently

available data to identify equity dimensions, analyse distributions of educational outcomes and provide disaggregated indicators for monitoring and evaluation of policies and programmes (Chapter 3). It has also called for improvements in data and management systems, particularly with respect to learning, in order to ensure better tracking of progress towards national and global equity-related targets in the future (Chapter 4).

The fourth gap concerns gender. This has been the primary dimension of equity being tracked with respect to access and, to some extent, learning in national education plans. While this attention to gender is welcome, the analysis of equity in national education plans has highlighted shortcomings with respect to other dimensions of equity that also require attention, such as poverty, location, ethnicity, language, religion and disability (Chapter 4).

The fifth gap relates to the need to establish common definitions, metrics and standards to ensure reliability and international comparability, as discussed in Chapter 4. While the handbook highlights the importance of nationally-relevant approaches to equity, using data sources that are contextually relevant and currently available, it also calls for establishing a common framework to achieve the international comparability needed for monitoring progress towards SDG 4 on education (Chapter 4).

Finally, in order to achieve equity in education, governments may choose to distribute educational inputs unequally, in ways that compensate for existing disadvantage (Chapters 2 and 5). The handbook has described ways to measure who benefits from government education expenditure, various funding strategies for redistribution of government financing, and finally how setting up national education accounts can help to monitor how funds are being spent in different countries to promote equitable outcomes (Chapter 5).

Drawing on the lessons learned from addressing these gaps, the handbook presents the following key messages:

1. A conceptual understanding of equity is needed to be clear about what we are measuring and why we are measuring equity in education.

2. There are a variety of equity metrics which could be used with education indicators, and many indicators have important properties which are relevant to the study of equity in education. However, some of these are not easy to communicate to policymakers and stakeholders in education, who therefore need guidance from statisticians and other experts so that all meaningful indicators can inform the national planning processes.

3. Impartiality approaches to equity have been the most widely adopted in national education plans. However, measures related to minimum standards and redistribution are also included in the SDG 4 indicator framework and can be used to track progress for sub-groups of the population.[75]

4. National education plans need to include a wider range of measures of equity in access and learning that go beyond gender. It is also important for plans to identify how intersecting disadvantages may hinder progress towards access and learning. Gender disparities are compounded, for example, by poverty, geographical location and disability.

75 One example for an indicator referring to minimum standards is SDG Indicator 4.1.1, the "proportion of children and young people (a) in Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics". Another example is SDG Indicator 4.6.1, the "proportion of a population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills". SDG Indicator 4.5.3, the "extent to which explicit formula-based policies reallocate education resources to disadvantaged populations", is an example for a measure referring to redistribution.

5. While there has been some progress towards including dimensions of equity for tracking progress in access to primary school and, to some extent, to secondary school, there is still a need for improved metrics of equity in learning at all levels of education.

6. It is necessary to measure equity in learning for children both in and out of school, since those out of school are likely to be at the greatest educational disadvantage.

7. A strong emphasis on measuring equity in access and learning from the early years is needed so that initial inequalities can be identified and targeted as early as possible.

8. Current progress in education cannot be tracked for the most disadvantaged groups unless there is a strong emphasis on improving educational management and information systems (EMIS) on access and learning, and to link these data to existing household surveys, which contain information about the socioeconomic and demographic characteristics of children, youth and adults.

9. Merely measuring equity in education is not enough. Systems of education also need to adopt equitable strategies involving redistribution of education funding, allocation of teachers and resources, as well as targeted approaches to raising learning standards for those at risk of being left behind.

Annex A. Major international learning

Learning Assessment	Geographic coverage	Domain	Data properties
Early Grade Reading Assessment (EGRA)	EGRA has been used in 70+ countries and has been adapted to 120+ languages by 30+ organizations.	Listening comprehension, oral language, alphabet knowledge, decoding, oral reading fluency, reading comprehension, phonological awareness, word recognition.	A number of different indicators from the EGRA assessment can be used as an outcome indicator. Alternatively, raw scores of subtasks or a composite index that utilises scores from across the subtasks can be developed. Benchmarks can also be developed to provide "the percentage of pupils achieving" at specified levels. These are not standardised across EGRA assessments.
Early Grade Math Assessment (EGMA)	EGMA has been used in 70+ countries and has been adapted to 120+ languages by 30+ organizations.	Addition and subtraction, number comparison, number patterns, procedural addition and subtraction knowledge, problem solving.	Number of correct items per minute, % correct attempted, % zero scores for each subtask.
Progress in International Reading Literacy Study (PIRLS)	PIRLS and pre-PIRLS have been administered in over 46 educational systems across Africa, Asia, Europe, and North and Central America.	Reading for literacy experience, reading to acquire and use information, processes of comprehension, reading behaviours and attitudes.	PIRLS generates scale scores which are developed for each pupil, allowing for a variety of analyses. In addition, four proficiency levels (plus a "did not meet benchmarks" category) are reported based on the benchmarks established from the scale scores.
People's Action for Learning (PAL) Network (including ASER and UWEZO)	ASER is administered in Pakistan and India. UWEZO is administered in Tanzania, Uganda and Kenya.	Literacy: Letters, words, paragraphs, story and comprehension. Numeracy: counting, number identification and operations of addition, subtraction, multiplication and division. The tests are set according to the Grade 2 level curriculum in each country.	Both raw scores and the percentage of pupils "passing" the Grade 2 test are reported. Raw scores refer to the % of items answered correctly. % of children who passed refers to the % of children who answered all Grade 2 items correctly.

assessments

	Education perspective	Comparability	Equity perspective
	As reading fluency measures the number of correct words read in a minute, it does not necessarily measure a pupil's ability to "read for meaning". Fluency as an indicator can overestimate reading abilities in pupils with increased phonological awareness, therefore it is more representative of the domain (reading) to develop a composite index that utilises scores from each subdomain.	EGRA is a language-based assessment, and therefore is not comparable across languages.	As EGRA assesses very early reading skills, it can be sensitive to ceiling and floor effects. Programme architects and policy makers should be cognisant that focussing too heavily on the percentage of pupils achieving benchmarks can lead teachers to focus on pupils closer to achieving the benchmark.
	EGMA assesses student's abilities to apply mathematical procedures and (less so) solve mathematical problems. Assessments in Nigeria and Tanzania show that even when children can recall rules in order to compute mathematical calculations, their ability to apply this knowledge to both everyday and new contexts often remains limited. Educators and policy makers need to remain aware of the broader learning needs of pupils.	EGMA is often adapted. As raw scores are used, EGMA is not comparable across tests.	As with EGRA, EGMA assesses very early mathematical skills. It can be sensitive to ceiling and floor effects.
	PIRLS is a high-quality assessment which focuses on measuring reading literacy across languages and contexts. This is useful to inform differences in reading achievement across contexts, but due to the focus on comparability, it may miss important nuances regarding the different learning trajectories of children by language and context.	PIRLS is comparable across participating countries and over time (with a break in comparability between 2000 and 2001).	Children who speak minority languages may be assessed in a second or third language. It is difficult to explicate the extent to which poor performance may be due to systematic equity concerns regarding specific language speaker groups versus the assessment being administered in a second language. While this issue cannot be fully resolved, it is important to remain cognisant of this tension. In addition, PIRLS is administered to in-school children, thus the most vulnerable children are unlikely to be included in the analysis.
	PAL Network tools provide information on what children know and can do at the highest level of generality in order to inform citizens on the question "are our children learning?". It provides less information on children's learning trajectories in order to inform more specific educational research questions. The tools and the administration process assume a learning trajectory for all children regarding which skills are acquired first. When a child can no longer answer questions, the assessment ceases. In most cases, this is likely to be a correct assumption, however in some cases children may develop knowledge and skills in a less predictable trajectory. This could lead to underestimating the performance of some children.	The assessments are comparable across countries to the extent that they compare the percentage of pupils who are achieving against the expectations of each country's curriculum. However, the performance of pupils is not directly comparable as the assessments are different across countries.	The quick assessment cycle of these assessments provides rapid information on whether children are learning and provides an accountability mechanism for parents and the community. This rapid turnaround is important for equity. The surveys assess children within households and therefore all children are included. However, as the tools are generally aimed at the lower levels of education, they are not able to capture differences between groups at the higher levels of education due to ceiling effects.

Learning Assessment	Geographic coverage	Domain	Data properties
Programme for International Student Assessment (PISA)	PISA is administered in most OECD member countries and is also extended to several countries across South and Central America, and Africa.	Scientific literacy: Understand scientific questions, explain scientific phenomena, use scientific evidence. Reading literacy: Retrieve texts and access them, interpret and integrate texts, and reflect and evaluate texts. Mathematical literacy: Quantify change and relationships, space and shape, uncertainty.	PISA generates scale scores which are developed for each pupil, allowing for a variety of analyses. In addition, 6 proficiency levels (per domain) are reported based on the benchmarks established from the scale scores.
The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)	SACMEQ consists of 16 ministries of education in Southern and Eastern Africa.	Reading: narrative, expository and documents. Mathematics: number, measurement and spatial data.	SACMEQ generates scale scores which are developed for each pupil, allowing for a variety of analyses. In addition, 8 proficiency levels (per domain) are reported based on the benchmarks established from the scale scores.
Programme d'analyse des systèmes éducatifs de La Confemen (PASEC)	PASEC has been administered in 13 countries in Francophone West Africa.	Language early primary: listening comprehension, familiarisation with written language and reading decoding, and reading comprehension. Language late primary: decoding isolated words and sentences, reading comprehension. Mathematics early primary: arithmetic, geometry, space and measurement. Mathematics late primary: arithmetic, measurement, and geometry and space.	PASEC generates scale scores for each student and there are competency levels for each domain. There are 4 language competency levels, with a "Sufficient" threshold at the cut-point between levels 2 and 3. There are 3 mathematics competency levels, with a "Sufficient" threshold at the cut-point between levels 1 and 2.
Pacific Islands Literacy and Numeracy Assessment (PILNA)	PILNA is administered across 14 countries in the Pacific region.	Literacy and numeracy in years 4 and 6.	PILNA generates scale scores for each student and there are competency levels (proficiency bands) for each domain. There are 9 proficiency bands ranging from level 0 to level 8.
Trends in International Mathematics and Science Study (TIMSS)	TIMMS is administered in 52 participating countries across a variety of regions.	Fourth grade mathematics: number, geometric shapes and measures, data display. Eighth grade mathematics: number, algebra, geometry, data and probability. Fourth grade science: life science, physical science and Earth science. Eighth grade science: biology, chemistry, physics and Earth science.	TIMSS generates scale scores for each student, and there are 4 proficiency levels (plus a "did not meet benchmarks" category).
Segundo y Tercer Estudio Regional Comparativo y Explicativo (SERCE/ TERCE)	13 countries across South America and the Caribbean participated in S/TERCE.	Third grade: mathematics, reading and writing. Sixth grade: mathematics, reading and writing, and natural sciences.	S/TERCE generates scale scores for each student, and places students into levels of performance ranging from level I (easiest tasks) to level IV (most difficult).

Sources:

ACER (2015). The Southern and Eastern Africa Consortium for Monitoring Educational Quality. Melbourne: ACER.

Brombacher, A., J. Bulat, S. King, E. Kochetkova, and L. Nordstrum (2005). National Assessment Survey of Learning Achievement at Grade 2: Results for Early Grade Reading and Mathematics in Zambia. Durham, NC: RTI International.

Gronmo, L. S., Lindquist, M., Arora, A., and I.V. Mullis (2013). TIMSS 2015 Mathematics Framework. Boston: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

OECD (n.d.). PISA 2012 Reading Framework. Paris: OECD.

OECD (n.d.). PISA - The OECD Programme for International Student Assessment: Building on PISA's Knowledge Base. Paris: OECD.

Education perspective	Comparability	Equity perspective
PISA aims to measure how much students approaching the end of compulsory education have acquired of the knowledge and skills essential for full participation in the knowledge economy. The assessments are not aligned to a national curriculum.	PISA is comparable across countries and over time.	PISA's target group is 16-year-olds. In many low- and middle-income countries the most vulnerable students are out of school. This has implications for measuring equity of learning.
SACMEQ measures proficiency in reading and mathematics across a number of countries. As with PIRLS, SACMEQ assesses literacy in the official languages of each country, but it may be that language rather than literacy is being assessed for some.	SACMEQ is designed to be comparable across countries and over time.	As with PIRLS and PISA, SACMEQ is administered to children in school in Grade 6 and does not include out-of-school children.
The late primary PASEC language assessment does not assess writing competencies, listening comprehension, oral communication skills or the tools specific to each language assessed (spelling, grammar and verb conjugations). These are important skills for language learners and thus this limitation should be kept in mind.	PASEC is designed to be comparable across countries and over time. However, there are some breaks in comparability due to efforts to improve comparability with other regional assessments.	As with PISA, SACMEQ and PIRLS, PASEC is administered to in-school children and does not include out-of-school children.
While experiencing some technical challenges, PILNA is moving towards high-quality assessment programme procedures. Countries were able to adopt translated test versions based on the individual language policies and the languages of instruction in each country, but the linguistic diversity within the Pacific region indicates that language, rather than literacy is being assessed for some.	PILNA is comparable both over time and between year levels, with achievement being reported on the same scale for years 4 and 6.	As with PISA, SACMEQ, PIRLS and PASEC, PILNA is administered to in-school children and does not include out-of-school children.
TIMSS is a high-quality assessment which focuses on measuring the ability of students to understand and act from a sound scientific basis and solve mathematical problems. It is not aligned to a single curriculum, though it does take the curricula of participating countries into account.	TIMSS is designed to be comparable across countries and over time.	As with PISA, SACMEQ, PIRLS, PASEC and PILNA, TIMSS is administered to in-school children.
S/TERCE is a high-quality assessment which focuses on measuring what a primary school student should learn and know, as defined by the current curricula in the participating countries. While this test informs about differences across contexts within the region, due to its focus on comparability it may miss important nuances regarding the different learning trajectories of children by language and context.	S/TERCE is designed to be comparable across countries and over time.	As with PISA, SACMEQ, PIRLS, PASEC, PILNA and TIMSS, S/TERCE is administered to in-school children.

Outhred, R. (2015). How to make the Sustainable Development Goal for Education Work. 7 May. Retrieved from The Economist Intelligence Unit.

PASEC (2015). PASEC 2014 Education System Performance in Francophone Sub-Saharan Africa: Competencies and Learning Factors in Primary Education. Dakar, Senegal: PASEC.

PASEC (2016). PASEC Education System Performance in CONFEMEN Countries. Competencies and Learning Factors in Primary School. GAML Steering Committee. PASEC.

Singh, A. (2015). How standard is a standard deviation? A cautionary note on using SDs to compare across impact evaluations in education. 13 January. http://blogs.worldbank.org/impactevaluations/how-standard-standard-deviation-cautionary-note-using-sds-compare-across-impact-evaluations

Uwezo (2014). Are Our Children Learning? Literacy and Numeracy Across East Africa 2013. Nairobi: Uwezo East Africa. https:///learningportal.iiep.unesco.org/en/notice/T1440949025

Annex B. 75 national education plans

Country	SDG region	Name of plan
Afghanistan	Central and Southern Asia	Education Sector Plan 2011-2013.
Albania	Europe and Northern America	National Education Strategy 2004 - 2015
Armenia	Northern Africa and Western Asia	Education Development State Programme of the Republic of Armenia 2011-2015.
Bangladesh	Central and Southern Asia	Sixth Five Year Plan FY2011-FY2015. Accelerating Growth and Reducing Poverty. (Part 1/3 Strategic Directions and Policy Framework)
Benin	Sub-Saharan Africa	Plan Décennal du Secteur de L´Éducation du Bénin. Troisième phase, couvrant la période 2013-2015.
Bhutan	Central and Southern Asia	Bhutan Education Blueprint 2014-2024. Rethinking Education
Bolivia	Latin America and the Caribbean	Plan Estratégico Institucional 2010-2014
Botswana	Sub-Saharan Africa	Education and Training Sector Strategic Plan (ETSSP 2015-2020). Republic of Botswana.
Brazil	Latin America and the Caribbean	Planejando a Próxima Década Conhecendo as 20 Metas do Plano Nacional de Educação
Burkina Faso	Sub-Saharan Africa	Programme Sectoriel de L´Éducation et de la Formation (PSEF) 2012-2021
Cambodia	Eastern and South-Eastern Asia	Education Strategic Plan. 2014-2018
Cameroon	Sub-Saharan Africa	Document de Stratégie du Secteur de l'Education et de la Formation
Central African Republic	Sub-Saharan Africa	Plan d'action du secteur de l'éducation de la République Centrafricaine pour 2013-2015
Chad	Sub-Saharan Africa	Stratégie intérimaire pour l'Education et l'alphabétisation
Congo	Sub-Saharan Africa	Stratégie Sectorielle de L'Education et de la Formation 2016-2025
Côte d'Ivoire	Sub-Saharan Africa	Plan d'Actions A Moyen Terme ŔPAMT Secteur Education/Formation. 2012-2014
Egypt	Northern Africa and Western Asia	Strategic Plan for Pre-University Education
El Salvador	Latin America and the Caribbean	Plan Nacional de Educacion 2021
Eritrea	Sub-Saharan Africa	Eritrea's Education Sector Development Plan 2013-2017
Ethiopia	Sub-Saharan Africa	Education Sector Development Programme V (ESDP V).
Gambia	Sub-Saharan Africa	Draft Education Sector Plan 2014-2022
Georgia	Northern Africa and Western Asia	Consolidated Education Strategy and Action Plan (2007-2011)
Ghana	Sub-Saharan Africa	Education Strategic Plan 2010 to 2020: I) Volume 1. Policies, Strategies, delivery finance. and II) Volume 2. Strategies and Work Programme.
Guatemala	Latin America and the Caribbean	Plan de Implementación Estratégica de Educación 2012-2016

Year	Global Partnership for Education	Weblink
2010	Yes	http://www.globalpartnership.org/content/islamic-republic-afghanistan-education-interim-plan-2011-13
2004	Yes	http://www.globalpartnership.org/content/republic-albania-national-education-strategy-2004-2015
2011		http://www.gdf.am/images/Education_Development_National_Programme_2011-2015_eng.pdf
Not dated	Yes	http://www.globalpartnership.org/content/sixth-five-year-plan-part-1-bangladesh
2013	Yes	http://www.globalpartnership.org/fr/content/benin-plan-decennal-developpement-secteur-education-actualise-phase-3-2013-2015
2014	Yes	http://www.globalpartnership.org/content/bhutan-education-blueprint-2014-2024
Not dated		http://medios.economiayfinanzas.gob.bo/MH/documentos/Planificacion/Plan_Estrat%C3%A9gico_Institucional.pdf
2015		http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/botswana_etssp_2015-2020.pdf
2014		http://pne.mec.gov.br/images/pdf/pne_conhecendo_20_metas.pdf
2012	Yes	https://www.globalpartnership.org/fr/content/plan-sectoriel-de-leducation-2012-2021-burkina-faso
2014	Yes	http://www.globalpartnership.org/content/education-sector-plan-cambodia
2013	Yes	http://www.globalpartnership.org/fr/content/plan-sectoriel-education-cameroun
2012	Yes	http://www.globalpartnership.org/fr/content/republique-centrafricaine-plan-daction-du-secteur-de-leducation-2013-2015
2012	Yes	http://www.globalpartnership.org/fr/download/file/fid/3202
2015	Yes	http://www.globalpartnership.org/fr/download/file/fid/50364
2014	Yes	http://www.globalpartnership.org/fr/republique-de-cote-divoire-plan-dactions-moyen-terme-pamt-secteur-educationformation
2014		http://www.unesco.org/education/edurights/media/docs/c33b72f4c03c58424c5ff258cc6aeaee0eb58de4.pdf
2004		http://www.oei.es/historico/quipu/salvador/plan2021_metasypoliticas.pdf
2013	Yes	http://www.globalpartnership.org/content/eritrea-education-sector-development-plan-2013-2017
2015	Yes	http://www.globalpartnership.org/library?lang[0]=en&lang[1]=fr&f[0]=field_document_type%3A1401&f[1]=field_country%3A132
2013	Yes	http://www.globalpartnership.org/content/republic-gambia-department-state-education-efa-proposal-fast-track-initiative
2006	Yes	http://www.globalpartnership.org/content/georgia-consolidated-education-strategy-and-action-plan-2007-2011
2012	Yes	http://www.globalpartnership.org/library?lang[0]=en&lang[1]=fr&f[0]=field_document_type%3A1401&f[1]=field_country%3A145
2012		http://infopublica.mineduc.gob.gt/mineduc/images/7/71/DIPLAN_Plan_de_Implementacion_Estrategica_de_Educacion_2012-2016.pdf

Country	SDG region	Name of plan
Guinea	Sub-Saharan Africa	Programme sectoriel de l'éducation. 2015-2017 [PSE] Guinée. Lettre d'endossement [LE]
Guinea-Bissau	Sub-Saharan Africa	Three-year Plan for the Development of Education: 2011 - 2013
Guyana	Latin America and the Caribbean	Education Sector Plans for 2014-2018. v1.
Haiti	Latin America and the Caribbean	La Stratégie Nationale D'Action pour L'Éducation pour Tous
Honduras	Latin America and the Caribbean	Plan Estratégico Institucional 2014 - 2018
Iran	Central and Southern Asia	Fundamental Reform Document of Education (FRDE), in the Islamic Republic of Iran
Kenya	Sub-Saharan Africa	National Education Sector Plan.
Kiribati	Oceania	I. Kiribati Development plan 2016-2019 II. Kiribati Education Sector Plan 2012-2015
Kosovo	Europe and Northern America	Kosovo Education Strategic Plan 2017-2021
Kyrgyzstan	Central and Southern Asia	Education Development Strategy of the Kyrgyz Republic for 2012-2020
Lao People's Democratic Republic	Eastern and South-Eastern Asia	Education Sector Development Plan (ESDP). (2011-2015)
Lebanon	Northern Africa and Western Asia	I. National Educational Strategy in Lebanon (Vision Document) II. National Educational Strategy in Lebanon (Strategic Plan Document)
Lesotho	Sub-Saharan Africa	Kingdom of Lesotho. Education Sector Plan
Liberia	Sub-Saharan Africa	The Education Sector Plan of Liberia- A Commitment to making a difference
Malawi	Sub-Saharan Africa	National Education Sector Plans for 2008-2017. A Statement.
Malaysia	Eastern and South-Eastern Asia	Malaysia Education Blue Print 2013-2025
Mali	Sub-Saharan Africa	Programme Intérimaire de Relance du Secteur de L'Education et de la Formation Professionnelle – Durée: 2 ans (2015-2016)
Mauritania	Sub-Saharan Africa	Programme National de Développement du Secteur Educatif 2011-2020 (PNDSE II). Plan D'Action Triennal (2012-2014)
Mexico	Latin America and the Caribbean	Programa Sectorial de Educación 2013-2018
Micronesia, Federated States of (Chuuk State)	Oceania	Chuuk State Strategic Plan for Education 2007-2012
Mongolia	Eastern and South-Eastern Asia	Master Plan to Develop Education of Mongolia
Mozambique	Sub-Saharan Africa	Education Strategic Plan. 2012-2016
Myanmar	Eastern and South-Eastern Asia	National Education Strategic Plan 2016-2021
Nepal	Central and Southern Asia	School Sector Reform Plan 2009-2015.
Nicaragua	Latin America and the Caribbean	Plan Estrategico de Educación 2011-2015
Niger	Sub-Saharan Africa	Programme Sectoriel de l'Education et de la Formation (2014-2024).
Nigeria (Kano State)	Sub-Saharan Africa	Education Sector Plan 2009-2018. Nigeria, Kano State.

Year	Global Partnership for Education	Weblink
2014	Yes	https://www.globalpartnership.org/fr/content/plan-sectoriel-de-leducation-guinee
2010	Yes	http://www.globalpartnership.org/content/republic-guinea-bissau-three-year-plan-development-education-2011-2013
2013	Yes	http://www.globalpartnership.org/content/education-sector-plan-guyana
2007	Yes	http://planipolis.iiep.unesco.org/fr/2007/strat%C3%A9gie-nationale-daction-deducation-pour-tous-snaept-4376
2014	Yes	http://www.globalpartnership.org/content/education-sector-plan-honduras
2011		http://www.dres.ir/safeschool/Downloads/FRDE.pdf
2014	Yes	http://www.globalpartnership.org/content/education-sector-plan-2013-2018-kenya
2016		https://dfat.gov.au/about-us/publications/Documents/kiribati-education-sector-strat-plan.pdf; http://www.mfed.gov.ki/sites/default/files/Kiribati%20Development%20Plan%202016%20-%2019.pdf
2016		http://www.kryeministri-ks.net/repository/docs/KOSOVO_EDUCATION_STRATEGIC_PLAN.pdf
2012	Yes	http://www.globalpartnership.org/country/kyrgyz-republic
2011	Yes	http://www.globalpartnership.org/content/education-sector-plan-lao-pdr
2007		http://www.laes.org/upload//editor_upload/file/Reports/strategicplanleb/Vision%20Document%20%20English.pdf; http://www.laes.org/upload//editor_upload/file/Reports/strategicplanleb/Strategic%20Plan.pdf
2005	Yes	http://www.globalpartnership.org/content/kingdom-lesotho-education-sector-plan-2005-2015
2009	Yes	http://www.globalpartnership.org/content/liberia-education-sector-plan
2008	Yes	http://www.globalpartnership.org/content/malawi-education-sector-plan
2013		http://www.moe.gov.my/images/dasar-kpm/articlefile_file_003108.pdf
2015	Yes	http://www.globalpartnership.org/fr/content/plan-interimaire-du-secteur-de-leducation-2015-2016-mali
2011	Yes	http://www.globalpartnership.org/fr/content/mauritanie-programme-national-de-developpement-du-secteur-educatif-2011-2020
2013		https://www.sep.gob.mx/work/models/sep1/Resource/4479/4/images/PROGRAMA_SECTORIAL_DE_EDUCACION_2013_2018_WEB.pdf
2007		http://www.paddle.usp.ac.fj/cgi-bin/paddle?e=d-010off-paddle--00-1--0---0-10-TX--6-------0-11l--11-en-50---20-png---10-3-1-000--0-0-11-0utfZz-8-00&a=file&d=chu001
2006	Yes	http://www.globalpartnership.org/content/master-plan-develop-education-mongolia-2006-2015
2013	Yes	http://www.globalpartnership.org/content/mozambique-education-strategic-plan-2012-2016
2016		https://firstrangoon.files.wordpress.com/2017/02/nesp-english.pdf
2009	Yes	http://www.globalpartnership.org/content/school-sector-reform-plan-2009-2015-nepal
2011	Yes	http://www.globalpartnership.org/content/nicaragua-plan-estrat%C3%A9gico-de-educaci%C3%B3n-2011-2015
2013	Yes	http://www.globalpartnership.org/fr/content/niger-plan-secteur-education-2014-2024
2008	Yes	http://www.globalpartnership.org/content/education-sector-plan-2009-2018-nigeria-kano-state

Country	SDG region	Name of plan
Pakistan (Sindh Province)	Central and Southern Asia	Sindh Education Sector Plan 2014-18
Palestine	Northern Africa and Western Asia	Palestine Education Development Strategic Plan EDSP 2014-2019
Papua New Guinea	Oceania	I. Achieving Universal Education for a Better Future. Universal Basic Education Plan. 2010-19 II. National Education Plan 2004-2015.
Philippines	Eastern and South-Eastern Asia	Philippine Development Plan 2011-2016. Midterm Update. With Revalidated Results Matrices.
Rwanda	Sub-Saharan Africa	Education Sector Strategic Plan for 2013-2018.
Samoa	Oceania	Samoa Education Sector Plan 2013-2018. Improved focus on Access to Education, and Training and Quality Learning Outcomes.
Sao Tome and Principe	Sub-Saharan Africa	Carta de Politica Educativa Sao Tome e Principe
Sierra Leone	Sub-Saharan Africa	Education Sector Plan 2014-2018. Learning to Succeed
Solomon Islands	Oceania	National Education Action Plan 2013-2015
Somalia (Somaliland)	Sub-Saharan Africa	Somaliland's Education Sector Strategic Plan 2012-2016
South Africa	Sub-Saharan Africa	Improving Education, Training and Innovation. National Development Plan: Vision for 2030, Chapter 9.
South Sudan	Sub-Saharan Africa	General Education Strategic Plan 2012-2017. Promoting Learning for All.
Sudan	Northern Africa and Western Asia	Interim Basic Education Strategy 2012
Swaziland	Sub-Saharan Africa	Education Sector Policy
Tajikistan	Central and Southern Asia	National Strategy of Education Development of the Republic of Tajikistan until 2020
Timor-Leste	Eastern and South-Eastern Asia	National Education Strategic Plan. 2011-2030
Tonga	Oceania	Tonga Education Policy Framework 2004-2019.
Uganda	Sub-Saharan Africa	Education Sector Plan 2010-2015
United Republic of Tanzania	Sub-Saharan Africa	Education Sector Development Programme (2008-17)
Uzbekistan	Central and Southern Asia	Education Sector Plan for 2013-2017
Vanuatu	Oceania	Vanuatu Education Sector Strategy 2007–2016
Yemen	Northern Africa and Western Asia	Yemen Education Sector Plan. Mid Term Results Framework 2013-2015
Zambia	Sub-Saharan Africa	Education Sector. National Implementation Framework III. 2011-2015
Zimbabwe	Sub-Saharan Africa	Education Sector Strategic Plan 2016-2020.

Year	Global Partnership for Education	Weblink
2013	Yes	http://www.globalpartnership.org/content/education-sector-plan-2014-2018-sindh-province-pakistan
2014		http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/palestine_education_development_strategic_plan_2014_2019.pdf
2009	Yes	http://www.globalpartnership.org/content/papua-new-guinea-universal-basic-education-plan-2010-2019
2014		http://www.neda.gov.ph/wp-content/uploads/2014/08/NEDA-PDP-2011-2016-Midterm-Update+Errata2.pdf; http://www.neda.gov.ph/wp-content/uploads/2013/10/pdprm2011-2016.pdf
2013	Yes	http://www.globalpartnership.org/content/rwanda-education-sector-strategic-plan-2010-2015
2013		http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/samoa_education_sector_plan_2013-2018.pdf
2012	Yes	http://www.globalpartnership.org/content/education-sector-plan-2012-2022-sao-tome-and-principe
2013	Yes	http://www.globalpartnership.org/content/sierra-leone-education-sector-plan
nership.ors3A.strTa8D9N01_Wrc5/tiitW0rTarue 5E	-plan-201_Wrcoad:cdO5P (s3A.strTa8D9N02g to Succeed)]TETE2E	