- Permanyer, I.; Smits, J. (2019) 'Inequality in human development across the globe'. Presented in Session E8. 8th Meeting of the Society for the Study of Economic Inequality (ECINEQ). Society for the Study of Economic Inequality. Paris (France), July 3-5. http://www.ecineq.org/ecineq_paris19/papers_EcineqPSE/paper_127.pdf

Inequality in human development across the globe

Iñaki Permanyer[1,*] and Jeroen Smits[2]

[1]Centre d'Estudis Demogràfics (a member of the CERCA Programme / Generalitat de Catalunya), Carrer de Ca n'Altayó, Edifici E-2, Campus de la UAB, 08193 Cerdanyola del Vallès, Spain. E-mail: ipermanyer@ced.uab.es. ORCID: 0000-0002-7051-5144.

[2]Global Data Lab, Institute for Management Research, Radboud University, Netherlands. E-mail: j.smits@fm.ru.nl. ORCID: 0000-0002-5802-975X

Abstract

The Human Development Index is the world's most famous indicator of the level of development of societies. A disadvantage of this index is however, that only national values are available, whereas within many countries huge subnational variation exists in income, health and education. Here we present the Subnational Human Development Index (SHDI), which shows within-country variation in human development and its dimension indices for over 1600 regions within 160 countries. The newly observed variation is particularly strong in low and middle developed countries (home to 70% of the world population) but less important in the most developed ones. While education disparities explain most of the SHDI inequality within low-developed countries, income differences are increasingly responsible for SHDI inequality within more highly developed countries. The new SHDI opens the possibility of studying global socio-economic change with unprecedented coverage and detail, increasing the ability of policy-makers to monitor the Sustainable Development Goals.

Introduction

Since 1990, the United Nations Development Program has reported on a yearly basis the values of its flagship indicator: the Human Development Index (*1*). This HDI – which indicates countries' combined achievements in education, health and standard of living – has become the key reference indicator to assess countries' socio-economic performance, in academic as well as policy-making circles (*2 – 6*). Despite its global success, the HDI has been criticized, among other issues, for the limited list of well-being dimensions it incorporates (*7 – 9*), the arbitrary choice of weights and aggregation function (*10 – 13*) and for not taking into consideration distributional issues in its conceptualization (*14 – 16*). In particular, the HDI has been criticized for its neglect of differences within countries (*17 – 19*). While there are multiple indications that subnational variation in human development can be substantial (*15 – 21*), until now it was not possible to study these differences for more than a handful of countries. Here, we document for the first time within-country variation in human development using comparable units of analysis *across the globe*. For that purpose, we use the new Subnational Human Development Index (SHDI) Database (*22*), which reports sub-national values of the HDI and its three subcomponents for more than 1600 regions within 160 countries that include above 99% of the world's population. Providing a ten times higher-resolution picture than was previously available, this source of data opens the possibility of studying global socio-economic change with unprecedented coverage and detail and increases the ability of policy-makers to monitor and achieve the Sustainable Development Goals. In this paper, we use the SHDI Database to report global trends in human development inequality since the year 2000 and investigate how much of that inequality can be attributable to differences occurring across or within countries.

Results

Figure 1 shows the distribution of the new index across the globe in 2015. We see that human development varies considerably, not only among but also within countries. The finely grained scale of the SHDI allows identifying clusters of high, middle and low human development that cut across national borders (e.g. the highly developed regions surrounding the Alps, the middle-developed coastal regions in front of the Gulf of Guinea, or the least developed regions in landlocked Sub-Saharan African countries). In other cases, one can observe clear geographic patterns within countries (e.g. the North-South divide in Belgium, Germany, Italy or Spain). Some countries exhibit huge regional variations (e.g. China, India, or Colombia) while others are quite homogeneous (e.g. Australia). Very often, the region where the capital city is located exhibits higher human development levels than the rest of the country.

[[[Figure 1: World Map]]]

SHDI inequality trends and decompositions

How has global inequality in SHDI and its three components evolved since the turn of the century? As shown in Figure 2, the Gini index for the world SHDI distribution has monotonically decreased from 0.14 in 2000 to 0.11 fifteen years later. This reduction of inequality has taken place against a backdrop of general improvements in human development across and within countries. Figure 2 also shows the trends in global inequality for the underlying health, education and income indices. While they are all decreasing, we observe substantial differences in their magnitudes and speed of decline. According to the Gini index, differences in the life expectancy index across world regions are smaller than differences in the education index. The former declined from 0.09 in 2000 to 0.08 in 2015, whereas the latter went from 0.19 to 0.14 in the same period. Somewhere in between, inequality in the income index went from 0.16 to 0.12. As the more unequal indices showed a faster decline, convergence is taking place towards the more equally distributed component (health).

One of the most interesting features of the SHDI and its subcomponents is that we can quantify the amount of inequality that can be attributed to differences occurring between and within countries. In Figure 3 we show the percentages of total inequality that can be attributed to within-country differences in the world as a whole and in the groups of 'Low', 'Medium' and 'High' developed countries in year 2015 (the results for 2000 are available in the Supplementary Materials Figure S1). These percentages can be seen as the amount of 'new' inequality that is revealed by using a subnational index. Most striking in Figure 3 is the very high contribution of within-country inequality to total inequality in the groups of countries at low and intermediate levels of development (where as much as 70 percent of the world population lives). In these groups of countries, about half of inequality in SHDI is within country inequality. Stated otherwise: when switching from national HDI to subnational SHDI, the amount of inequality we observe in low and middle developed countries doubles. Among the high developed countries, within-country inequality is less important and most SHDI inequality is explained by differences between countries. A rather similar pattern is observed when inspecting the three subcomponents of the SHDI separately (see Fig 3). Once again, the percentage of total inequality in education, health and income that can be attributed to within-country inequality is relatively small for the group of high developed countries (contributions hovering around 10% and 15%), but is much larger – and generally increasing – for low and, particularly, for medium developed countries (with contributions often approaching or even surpassing the threshold of 50%). In those regions, the newly found education, health and income differences within countries carry about the same weight as the (already well-investigated) differences across countries.

[[[Figure 2: Global inequality trends]]]

[[[Figure 3: % contribution of within country inequality]]]

Moving forward, one might wonder how much of the newly discovered HDI inequality within countries can be attributed to differences in education, health and income. Figure 4 plots for the groups of low, medium and high-developed countries, the contribution of the health (H), education (E) and income (I) subcomponents to SHDI inequality within countries (the Supplementary Materials section explains how to perform such decompositions). The upper, middle and lower panels show the results. Each dot in these ternary plots represents the three contributions for a specific country-year combination. The closer a dot is to a given vertex, the more important the corresponding subcomponent (H, E or I) is in explaining SHDI inequality within countries.

In the least developed countries, we observe that most observations (69%) are located near the education vertex (E). This means that in those countries, within-country SHDI inequality is mostly due to variation in education. At the intermediate level of development, education is still most important in explaining within-country variation, but the influence of income is growing (52% of points have E as its closest vertex, and 33% of points are closest to the I vertex). In the high developed countries, income surpasses education as the most important explanatory factor for within country SHDI variation (50% of the points have I as its nearest vertex in the lower panel). For all three groups of countries, differences in the health component tend to be least important. Its influence is relatively strongest at the intermediate level of development (with 15% of points nearest to H) and weakest in the most developed countries (with only 5% nearest to H).

[[[Figure 4: triangles]]]

Discussion

By introducing subnational variation, we obtain a richer and more comprehensive picture of the world's distribution of human development. The findings reported in this paper suggest that this distribution is characterized by (i) global inequality declines in the SHDI and its three subcomponents, (ii) large contributions of within-country inequality to overall inequality for the low and middle developed countries, and (iii) relatively small differences across subnational regions within highly developed countries. We find that variation in human development within countries can be as large as variation across countries in vast regions of the world. National-level aggregates such as the United Nations' Human Development Index conceal huge inequalities within low- and middle-income countries (home of 70% of the global population). While between-country inequality is still an important source of global variation in human development, within-country variability has gained increasing prominence during the last decades.

These patterns complement recent findings reported in global inequality studies (*23 – 28*). They are opposite to the long-established trends of ever-expanding between-country inequality, that date back as far as the early stages of the Industrial Revolution (*26, 28 – 30*). The shifting composition of human development inequality across and within countries implies that national location – while still of great importance – is losing prominence as a determining factor of individuals' well-being. The implications of this change, which are multifarious and far-reaching, will shape the 21st century agenda of social scientists and policy-makers concerned with global distributive justice.

Some limitations of this study should be highlighted. First, several data points in the SHDI Dataset are not based on 'real' observations, but have been estimated using different techniques. On the one hand, missing sub-national observations for the health and standard of living components have been estimated on the basis of under 5 mortality rates and the International Wealth Index (IWI), respectively (details given in the Materials and Methods section). On the other hand, missing data points were filled using simple interpolation and extrapolation techniques. The analyses performed in (*22*) demonstrate that the measurement error that entails from such estimation techniques is small (the fit of the models using proxy variables to estimate the health and standard of living components is very good – with R^2s around 89% and 83%, respectively – and the estimated relative errors for inter and extrapolation are quite small). In addition, the fact that our sub-national estimates are re-scaled so that their population weighted average coincides with the official national level guarantees the consistency of the dataset (see the Materials and Methods section). Lastly, while the measurement errors could somewhat affect our SHDI estimates for some specific sub-national regions, it is unlikely that they distort the *overall trends* reported in this article.

Second, the results presented in this paper uncover sub-national variation in human development at the level of the main administrative units of each country, but might fall short of explaining the variation that could remain within those units. Some exploratory analysis performed for the cases of Mexico (*20*) and Brazil (*31*) with a municipal-level HDI suggest that variations within countries' main administrative units can be potentially large. While greater geographical detail (e.g. municipalities (*20*) or small-scale grids (*32, 33*) can be obtained for specific (groups of) countries, this is obtained at the cost of losing comparability, thus missing the global picture. In addition, concepts like regional life expectancy, educational performance or standard of living can lose their meaning if the region becomes too small. The major strength of this study is that – for the first time – subnational variations in the HDI and its subcomponents are documented in a strictly comparable way for virtually countries of our world.

Like *any* composite index, both the HDI and its subnational version discussed here have their shortcomings when it comes to measure countries' or regions' socio-economic development levels. Both fail to take into consideration important dimensions of well-being (e.g. environmental sustainability indicators), and the way in which they aggregate information

(i.e. the choice of dimensional weights and aggregation functions) is somewhat arbitrary. Despite these well-known limitations, the HDI has played a fundamental role during the last decades to raise awareness that 'development is much more than economic growth' (*4*). Analogously, the new SHDI can be an extremely useful tool in the future to remind that, beyond efficient (i.e. aggregate) outcomes, whether or not these outcomes are equally distributed across the population is also a fundamental development goal.

There is a huge literature on within- and between-country inequalities in income and non-income dimensions of welfare. For example, many other studies have investigated sub-national variation in education (*27, 29, 32*), health (*23, 24*), or income (*25, 26, 28, 33*) across the world or in some of its regions. Our study differs from the previous ones in documenting global trends in education, health and standard of living *simultaneously* – thus offering a unique opportunity to investigate how these factors have *jointly* evolved and influenced each other over time. The data on the SHDI and its components is freely accessible in https://hdi.globaldatalab.org/. This provides researchers worldwide with high-detail contextual variables that can be used to improve our understanding in wide-ranging areas of the social sciences (including, but not limited to, the study of family formation and fertility behavior, migration, health and mortality, epidemiology, cultural / ideational / normative change, religion, socio-economic change, or environmental sustainability). In the policy-making arena, the subnational SHDI is very pertinent for the global development agenda, which is committed to 'improve health and education', 'economic growth' and 'reduce poverty and inequalities between and within countries' (i.e. Sustainable Development Goals #1, #3, #4, #8 and #10). With the SHDI and its indices, problem areas with respect to these SDGs can be detected so that resources can be directed to the places where they are mostly needed. Moving beyond country-level averages, the SHDI has the potential to bring equity concerns to the fore, and to serve as a key tool to articulate national and international development policies into a coherent whole.

Materials and Methods

All the details on how the SHDI Database has been constructed can be found in (*22*), which can be downloaded from https://hdi.globaldatalab.org/areadata/about-shdi/. In this section, we will summarize the key points and refer the interested reader to that paper for further details.

Methodologically, the SHDI is a translation of UNDP's official HDI (hdr.undp.org) to the subnational level. As such, it is an average of the subnational values of three basic dimensions: 'Education', 'Health' and 'Standard of living'. The specific indicators used in their definition include 'Mean years of schooling of adults aged 25+', 'Expected years of schooling of children aged 6', 'Life expectancy at birth' and 'Gross National Income per

capita (PPP, 2011 US$)'. These indices are measured using a variety of data sources, ranging from censuses to socio-economic and demographic household surveys. More specifically, the Subnational Human Development Index Database was created on the basis of three data sources: (i) statistical offices, including Eurostat, the statistical office of the European Union, (ii) the Area Database of the Global Data Lab, GDL-AD, (www.globaldatalab.org/areadata) and (iii) the HDI database of the United Nations Development Program (UNDP, https://hdr.undp.org/data).

In countries where the required indicators are not available at a subnational level, we estimate them using available information on related indicators. More specifically, subnational life expectancy and GNI per capita have been estimated from information on under 5 mortality (U5M) and the International Wealth Index (IWI, see (*34*)), respectively. In both cases, separate regression models were constructed that explained the variation of our variable of interest (i.e. life expectancy and GNI per capita) at the national level (obtained from the UNDP database) on the basis of national U5M and IWI scores derived from GDL-AD. We compared models with linear and nonlinear effects based on their adjusted R^2 and chose the model with the best fit. In the case of life expectancy, the selected model explained 89% of the variance, and in the case of GNI per capita, 83%.

To ensure consistency with national HDI values, the subnational indicators are constructed in such a way that their averages correspond to the official values published by UNDP. Using the later as reference points, we fill missing country-year information using interpolation and extrapolation techniques. In (*22*), details are given on (i) the extent of inter- and extrapolation performed on the SHDI Dataset, and (ii) the estimation of the relative error derived from the adoption of these techniques. Overall, the share of high-quality estimates (i.e. real observations, or short term inter- or extrapolations) increases over time, and the size of the relative errors tends to be smaller than the ones reported in analogous studies, e.g. (*35*). With the implementation of a consistent methodology, we obtain a simple measure spanning from 2000 to 2015 that is internationally comparable across more than 1600 regions in 160 countries, covering over 99% of the world population.

To compute the Subnational Human Development Index, we first estimate the education, health and standard of living subcomponents (E_i, H_i, S_i) and scale them between 0 and 1 (see Supplementary Materials section for details). Mimicking the most recent definition of UNDP's HDI, the Subnational Human Development Index for each subnational area 'i' is defined as

$$SHDI_i^m = \sqrt[3]{H_i E_i S_i} \qquad [1]$$

This is the geometric mean of the three sub-components (the superscript 'm' stands for 'multiplicative'). In this way, one avoids perfect substitutability between the health, education and standard of living dimensions and penalizes those regions with very unequal achievements across components. Like the original HDI, the SHDI takes values between 0 and 1 (the former is reached whenever one of the three components attains the lowest possible level of 0 and the latter when all three components attain the maximal level of 1). While not attempted in this paper, it is also possible to construct the additive version of the SHDI with the information available in the SHDI Database.

To measure the extent of global inequality in the SHDI and its sub-components, we have used two well-known inequality measures: the Gini index and the Mean Log Deviation. These measures have been chosen for their popularity and their well-known decomposability properties. The Mean Log Deviation allows decomposing global SHDI inequality in its within-country and between country components, and the Gini index can be used to assess how much of the SHDI inequality in a given country can be attributable to the education, health or standard of living components (details and formal definitions are given in the Supplementary Materials section).

References and Notes

1. UNDP (United Nations Development Programme)–Human Development Report Office. *Human Development Reports 1990, 1991, 1993, 1995, 2009, 2010* (Oxford University Press 1990 through 2005; and Palgrave Macmillan since 2006).
2. J. Bongaarts, S. Watkins, Social interactions and contemporary fertility transitions, *Population and Development Review.* **22(4),** 639-682 (1996).
3. M. Myrskylä, H. P. Kohler, F. Billari, Advances in development reverse fertility declines. *Nature*, **460**, 741-743 (2009).
4. J. Klugman, F. Rodríguez, H. Choi, H, The HDI 2010: New Controversies, Old Critiques. *Journal of Economic Inequality*, **9(2)**, 249-288 (2011).
5. F. Bray, A. Jemal, N. Grey, J. Ferlay, D. Forman, "Global cancer transitions according to the Human Development Index (2008-2030): a population-based study", *The Lancet Oncology*, **13(8)**, 790-801 (2012).
6. H. Chang, et al., Preventing preterm births: analysis of trends and potential reductions with interventions in 39 countries with very high human development index. *The Lancet*, **381(9862)**, 223-234 (2013).
7. S. Anand, A. Sen, Human Development Index: Methodology and measurement. Human Development Report Office Occasional Paper 12, UNDP, New York (1992).
8. E. Neumayer, The human development index and sustainability — a constructive proposal. *Ecological Economics*, **39** (1), 101-114 (2001).

9. G. Ranis, F. Stewart, E. Samman, Human Development: Beyond the Human Development Index. *Journal of Human Development*, **7**(3), 323-358 (2007).
10. L. Cherchye, E. Ooghe, T. Van Puyenbroeck, Robust human development rankings. *Journal of Economic Inequality*, **6**(4), 287-321 (2008).
11. J. Foster, M. McGillivray, S. Seth, Composite Indices: Rank Robustness, Statistical Association, and Redundancy. *Econometric Reviews*, **32**(1), 35-56 (2013).
12. I. Permanyer, Uncertainty and robustness in composite indices rankings. *Oxford Economic Papers*, **64** (1), 57-79 (2012).
13. M. Ravallion, Troubling tradeoffs in the Human Development Index. *Journal of Development Economics*, **99**(2), 201-209 (2012).
14. A. Sagar, A. Najam, The human development index: a critical review. *Ecological Economics*, **25**, 249-264 (1998).
15. J. Foster, L. Lopez-Calva, M. Szekely, Measuring the distribution of Human Development: methodology and an application to Mexico. *Journal of Human Development*, **6(1)**, 5-25, (2005).
16. M. Grimm, K. Harttgen, S. Klasen, M. Misselhorn, A Human Development Index by Income Groups. *World Development*, **36(12)**, 2527-2546 (2008).
17. M. Grimm, K. Harttgen S. Klasen, M. Misselhorn, T. Munzi, T. Smeeding, Inequality in Human Development: An Empirical Assessment of 32 Countries. *Social Indicators Research*, 97, 191-211.
18. K. Harttgen, S. Klasen, A Human Development Index by Internal Migration Status. *Journal of Human Development and Capabilities.* **12(3)**, 393-424 (2011a).
19. K. Harttgen, S. Klasen, A Household-based Human Development Index. *World Development.* **40(5)**, 878-899 (2011b).
20. Permanyer, I. (2013) "Using Census Data to Explore the Spatial Distribution of Human Development", *World Development.* **46**, 1-13 (2013)
21. I. Permanyer, J. García, A. Esteve, R. McCaa, Human Development Index-like Small Area Estimates for Africa compute from IPUMS-International integrated census microdata. *Journal of Human Development and Capabilities.* **16(2)**, 245-271 (2015).
22. J. Smits, I. Permanyer, The Subnational Human Development Database, *Global Data Lab Working Paper* 18-101.
23. J. Smits, C. Monden, Length of life inequality around the globe. *Social Science & Medicine.* **68**, 1114-1123 (2009).
24. R. D. Edwards, Changes in world inequality in length of life: 1970–2000. *Population and Development Review.* **37(3)**, 499-528 (2011).
25. S. Anand, P. Segal, The global distribution of income. In Atkinson, A. and Bourguignon, F. (Eds.) *Handbook of Income Distribution, Volume 2A* (Chapter 11), North Holland, 2015).
26. B. Milanovic, *Global inequality. A new approach for the age of globalization*. (Harvard University Press, 2016).
27. V. Jordá, J. Alonso, New estimates on educational attainment using a continuous approach (1970-2010). *World Development.* **90**, 281-293, (2016).

28. F. Bourguignon, C. Morrisson, Inequality among world citizens: 1820-1992. *American Economic Review*. **92(4)**, 727-744 (2002).
29. C. Morrisson, F. Murtin, The Kuznets curve of human capital inequality: 1870–2010. *The Journal of Economic Inequality.* **11(3)**, 283-301 (2013).
30. J. Riley, *Rising life expectancy: A global history*. (Cambridge University Press, 2001).
31. UNDP, Índice de Desenvolvimento Humano Municipal Brasileiro. Brasília, Ipea, FJP (2013).
32. N. Graetz, et al., Mapping local variation in educational attainment across Africa, *Nature*, **555**, (2018).
33. W. D. Nordhaus, X. Chen, Global Gridded Geographically Based Economic Data (G-Econ), Version 4. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC), (2016). http://doi.org/10.7927/H42V2D1C. Accessed 11/12/2018.
34. J. Smits, R. Steendijk, The International Wealth Index (IWI). *Social Indicators Research*, **122**(1), 65-85 (2015).
35. M. Kummu, M. Taka, J.H.A. Guillaume, Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015. *Scientific Data*. 5:180004, 1-15 (2018).
36. J. Smits, Sub-national development indicators for research and policy-making. *GDL Working paper* 16-101 (2016).
37. UNDP, Technical notes. Human Development Report 2016. United Nations Development Programme (2016).
38. A. Villar, *Lectures on Inequality, Poverty and Welfare*. Lecture Notes in Economics and Mathematical Systems 685, Springer (2017).
39. R. Lerman, S. Yitzhaki, Income inequality effects by income sources: a new approach and applications to the United States. *The Review of Economics and Statistics*, **67** (1), 151-156 (1985).
40. A. Shorrocks, Inequality decomposition by factor components. *Econometrica*, **50**(1), 193-211 (1982).

Acknowledgments

The research leading to these results has received funding from the European Research Council (ERC-2014-StG-637768, EQUALIZE project) and the Spanish Ministry of Economy and Competitiveness “Ramón y Cajal” Research Grant Program (RYC-2013-14196).

Figures and Tables

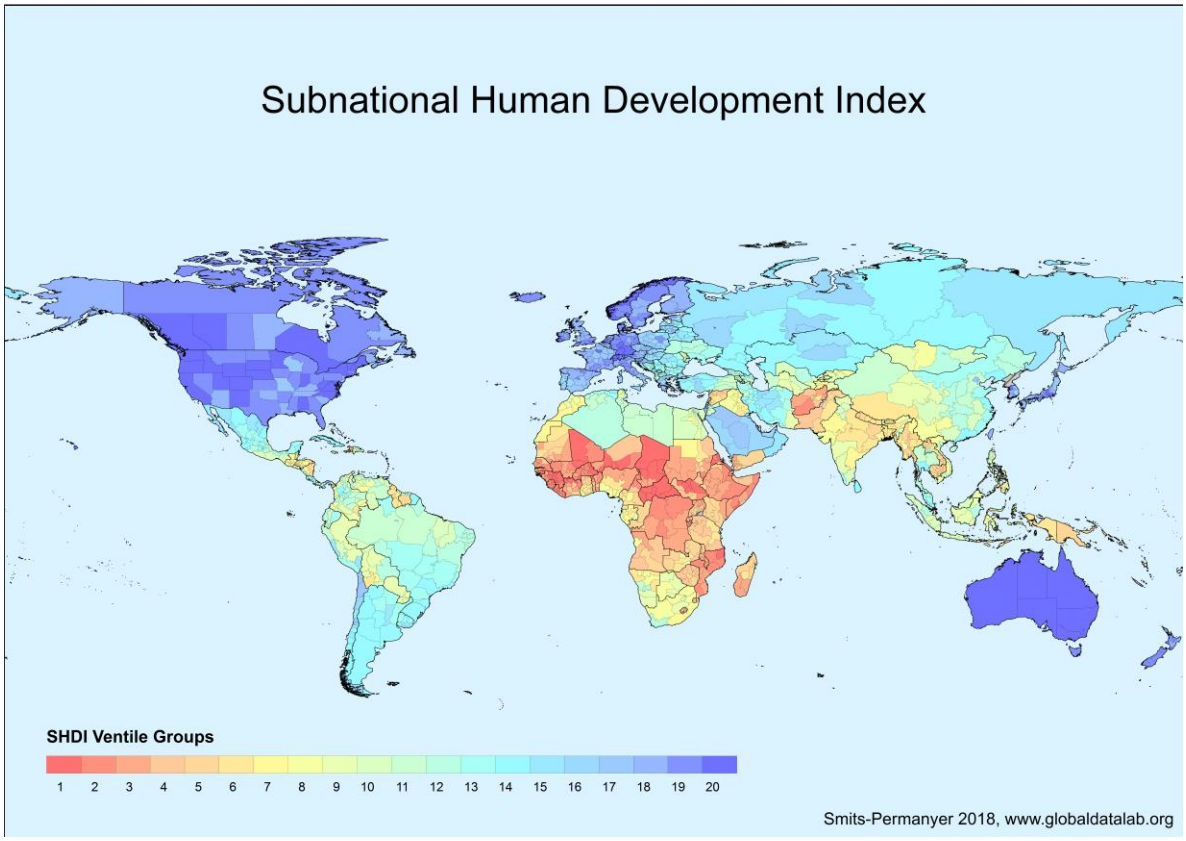

Figure 1. Global quantile distribution of the Subnational Human Development index in 2015. Source: Authors' elaboration based on information from the Global Data Lab.

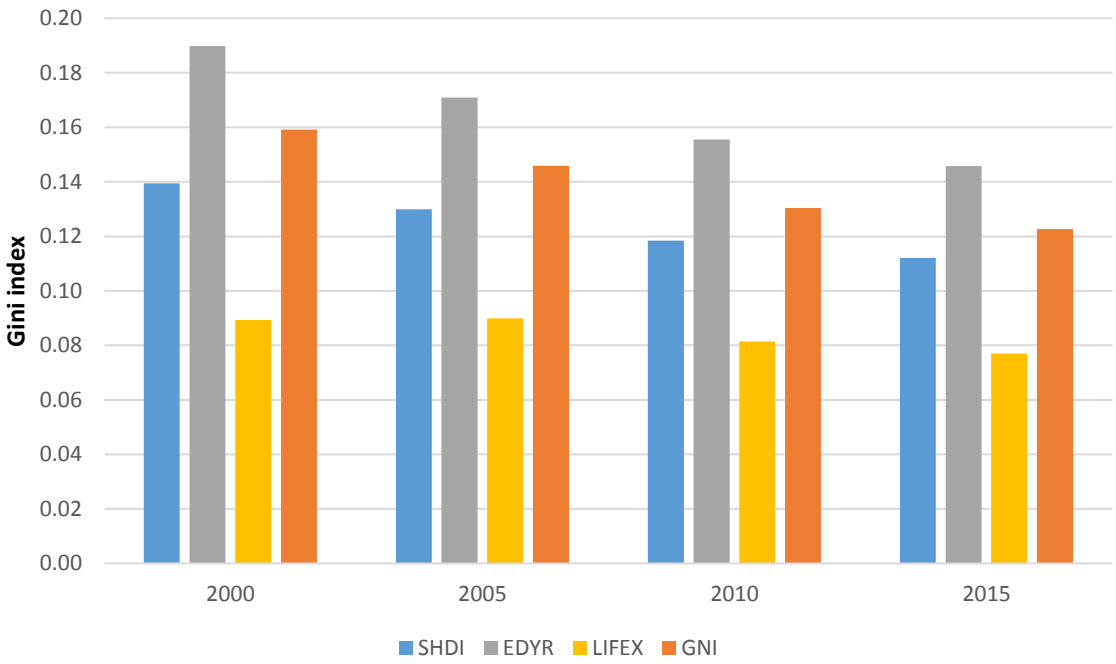

Figure 2. Global inequality trends across +1600 subnational regions (as measured with the Gini index) for the SHDI and its education, health and income sub-indices in years 2000, 2005, 2010 and 2015. Source: Authors' elaboration based on information from the Global Data Lab.

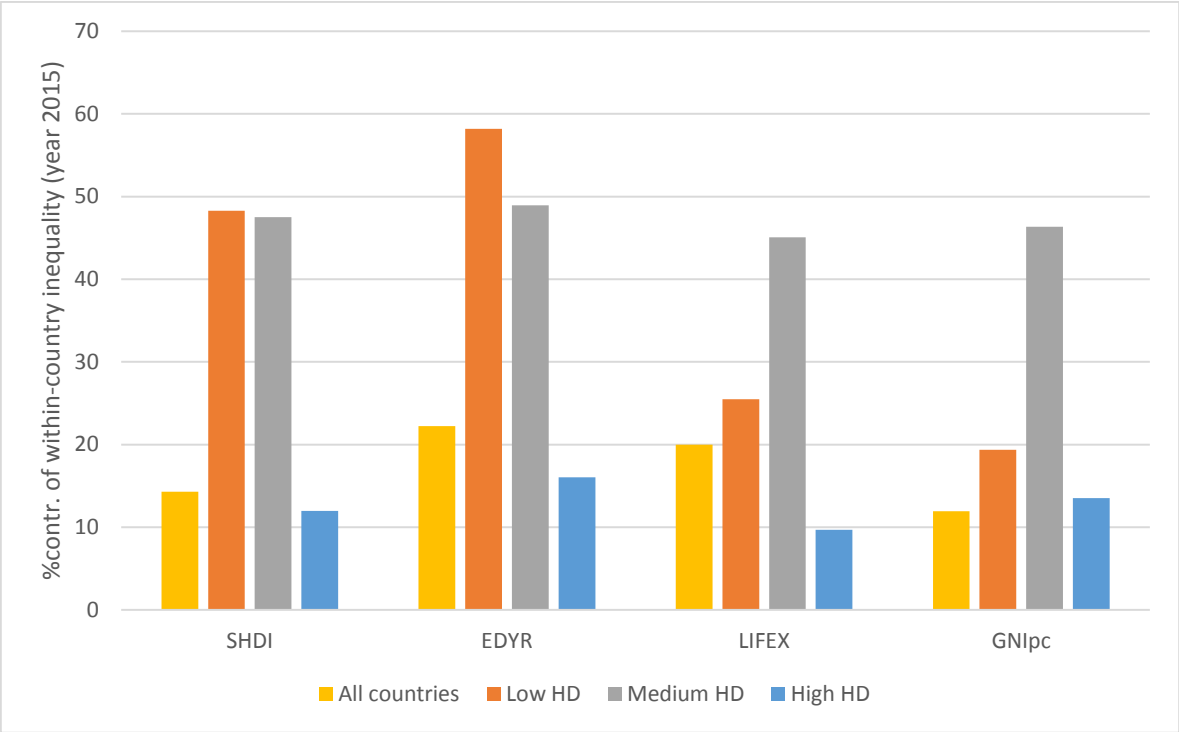

Figure 3. Percent contribution of within country inequality to total inequality for the SHDI and its three subcomponents in different groups of countries (year 2015). To measure the within-country contribution we have used the Mean Log Deviation as inequality measure (which is additively decomposable and is very highly correlated with the values of the Gini index – see Supplementary Information section). To define whether a country belongs to the group of Low, Medium or Highly developed countries we look at its HDI in year 2008 (the middle of the observation period). If such HDI is below (resp. above) 0.55 (resp. 0.7) the country belongs to the set of Low (resp. High) developed countries. Whenever it falls between 0.55 and 0.7, it is a Medium developed country. The shares of the world population in Low, Medium and High developed countries are 20%, 50% and 30%, respectively. The 50% horizontal line indicates the threshold above which the within-country component of inequality contributes *more* to total inequality than the between-country one. Source: Authors' elaboration based on information from the Global Data Lab.

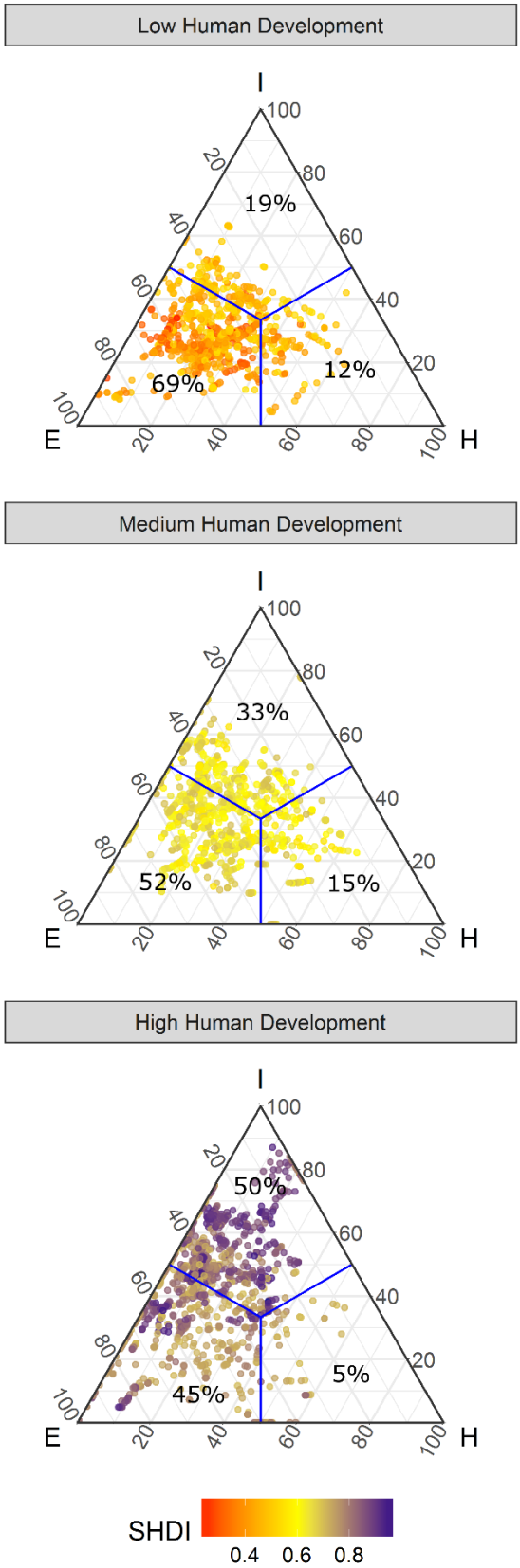
Low Human Development
I
E
H
19%
69%
12%
Medium Human Development
33%
52%
15%
High Human Development
50%
45%
5%
SHDI
0.4
0.6
0.8

Figure 4: Contribution of Education (*E*), Health (*H*) and Income (*I*) dimensions to SHDI inequality within countries by level of development. The upper, middle and lower panels show the contributions for the groups of Low, Medium and High developed countries, respectively. Explanations of how to estimate the contribution of each indicator are given in the Supplementary Materials. Ternary plots show the contribution of the three components to SHDI inequality within each country: the position of each dot with respect to the vertices helps understanding how important the education, health and income subcomponents are. The closer a dot is to a given vertex (say, *E*, *H* or *I*), the more important the corresponding subcomponent is. A hypothetical point in the middle of the triangle would represent a case where the three subcomponents contributed equally to observed SHDI inequality in a specific country. Another point near the top of the triangle would represent a case where most of the SHDI inequality would be explained by the income component. The numbers inside the triangles indicate the percentage of observations that are closest to the *E*, *H* or *I* vertices. Source: Authors' elaboration based on information from the Global Data Lab.

Supplementary Materials

The SHDI Database used in this paper contains subnational data for 1621 regions in 160 countries, that cover all areas and development levels across the globe. This is a big step forwards, especially for low and middle income countries (LMICs). Until recently there was hardly any subnational data for these countries available. The subnational data for LMICs used in the current research is to a large extent derived from the Global Data Lab (www.globaldatalab.org), which recently has released its GDL Area Database from where a broad range of subnational development indicators can be downloaded for over 120 LMICs (*36*). Much of the data for LMICs that we use to construct the SHDI was obtained from this GDL Area Database. For high-income countries and some middle-income countries, the data was derived from statistical offices, including Eurostat, the statistical office of the European Union. The SHDI was constructed in such a way that its national values are equal to the values of the official HDI of the United Nations Development Program (UNDP). The national data used for making this connection were derived from the UNDP HDI Database (hdr.undp.org/data).

Dimension indices

To compute the SHDI and its three subcomponents on the basis of the data we assembled, exactly the same methodology is used as the one employed by UNDP in the construction of the national HDI (*37*). The education sub-index is computed as $\left((EYS/18)\cdot(MYS/15)\right)^{1/2}$, where EYS and MYS are the expected and mean years of schooling, respectively. The health

sub-index is computed as $(e_0 - 20)/(85 - 20)$, where e_0 is the life expectancy at birth. Thirdly, the income sub-index is computed as $\big(log(GNIpc) - log(100)\big)/\big(log(75000) - log(100)\big)$, where $GNIpc$ is the Gross National Income per capita (2011 PPP $). Lastly, as shown in the main text (see equation [1]), the SHDI is computed as the geometric mean of the education, health and income sub-indices (the use of the arithmetic mean to generate an additive SHDI does not alter the main findings of the paper). Further details on materials and methods can be found in (*22*).

Measuring inequality

The extent of inequality in a given distribution can be measured using many different indices (see, for instance, (*38*)). In this paper, we use two of the most popular inequality measures: the Gini index (G) and the Mean Log Deviation (L). For a given distribution $\boldsymbol{x} = (x_1, \cdots, x_n)$ with $x_i \geq 0$ for all i, the Gini index is defined as

$$G(x_1, \cdots, x_n) = \frac{\sum_i \sum_j |x_i - x_j|}{2n^2\mu} \quad [EQ1]$$

where μ is the mean of the $\boldsymbol{x}$ distribution. The Gini index takes values between 0 and 1. When there is no variation and all observations are the same, $G = 0$. In the extreme case where all observations except one are zero (i.e. 'one individual owes everything and the others are left empty-handed'), $G = 1$.

The Mean Log Deviation is defined as

$$L(\boldsymbol{x}) = \frac{1}{n}\sum_i log\left(\frac{\mu}{x_i}\right) \quad [EQ2]$$

Like the Gini index, L takes a value of 0 whenever there is no variation and all observations coincide. Unlike the Gini index, L is unbounded: the higher its values, the higher the corresponding level of inequality.

Decomposing inequality

In order to decompose overall inequality in its within- and between-country components, we use the Mean Log Deviation, which belongs to the class of Generalized Entropy Measures. The reason why we have not used the Gini index in the decomposition is that, unlike L, the former cannot be neatly broken down into a within- and between-country component. Yet, the correlation between the values of the Gini index and the Mean Log Deviation for the distributions analyzed in this paper is extremely high (i.e. around 0.99).

The Mean Log Deviation is well known for its additive decomposability property. Assuming our distribution is split across several groups (as is the case with the SHDI distribution, where subnational regions are included within countries), overall inequality can be neatly decomposed as

$$L(\boldsymbol{x}) = L^B(\boldsymbol{x}) + L^W(\boldsymbol{x}) \quad [EQ3]$$

where L^B is the inequality between countries and L^W is the inequality within countries. The between-country component is obtained assuming that each subnational attainment is the same as the average attainment of the country to which the region belongs and the within-country component is a population-weighted average of the subnational inequality within countries. Formally,

$$L^B(\boldsymbol{x}) = \sum_{j=1}^{J} p_j log\left(\frac{\mu}{\mu_j}\right) \quad [EQ4]$$

$$L^W(\boldsymbol{x}) = \sum_{j=1}^{J} p_j L_j \quad [EQ5]$$

where J is the number of countries and p_j, μ_j and L_j are the population share, the mean and the Mean Log Deviation observed in country j, respectively. The percent contribution of within-country inequality to total inequality reported in Figure 3 is simply the quantity $100L^W(\boldsymbol{x})/L(\boldsymbol{x})$, which can be interpreted as the amount of 'new' inequality that is revealed by the subnational SHDI.

Contribution of the three components to within country SHDI inequality

The SHDI is an average of its three basic components. To reduce inequality in SHDI most effectively, it is important to know the exact contribution of the three components to the observed SHDI inequality levels. To estimate these contributions, methods can be used that have been developed in the economics literature to study the contribution of different income sources to total income inequality (e.g.: (*39*) or (*40*)). For these methods to work in our framework we need to use the additive version of the Subnational Human Development Index (which is very highly correlated with its multiplicative version). For a given subnational region $'i'$, let Y_i, H_i, E_i and S_i be the corresponding subnational human development, health, education and standard of living indices. In case of additive human development indices we have that

$$Y_i = \frac{H_i}{3} + \frac{E_i}{3} + \frac{S_i}{3} \quad [EQ6]$$

The distribution of subnational human development, health, education and standard of living indices will be denoted as *Y*, *H*, *E* and *S* respectively. Following (*40*), if the human development distribution is ordered so that $Y_1 \leq Y_2 \leq \cdots \leq Y_n$, then the corresponding Gini index can be written as

$$G(Y) = \frac{2}{n^2 \mu_y} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) Y_i \quad [EQ7]$$

where n is the number of regions andμ_y is the mean of the human development distribution. Plugging equation [EQ6] into equation [EQ7] we obtain

$$G(Y) = \frac{2}{3n^2 \mu_y} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) (H_i + E_i + S_i) = \frac{\mu_h}{\mu_y} \bar{G}(H) + \frac{\mu_e}{\mu_y} \bar{G}(E) + \frac{\mu_s}{\mu_y} \bar{G}(S) \quad [EQ8]$$

where μ_h, μ_e andμ_s are the means of the health, education and income distributions, and

$$\left.\begin{aligned} \bar{G}(H) &= \frac{2}{3n^2 \mu_h} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) H_i \\ \bar{G}(E) &= \frac{2}{3n^2 \mu_e} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) E_i \\ \bar{G}(S) &= \frac{2}{3n^2 \mu_s} \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right) S_i \end{aligned}\right\} \quad [EQ9]$$

which are known as the pseudo-Ginis for factors H, E and S respectively (see (*39,40*)). This way of rewriting the Gini coefficient provides a natural additive decomposition rule where the contributions of the different sources are clearly established.

The different inequality decomposition techniques used in this paper have been implemented using STATA.

Percent contribution of within-country inequality to total inequality in 2000.

In order to investigate how the contribution of within-country inequality to total inequality has evolved over time, in Figure S1 we show the analogue of Figure 3 (which shows the values corresponding to year 2015) for the year 2000. The contribution of within-country

inequality to total inequality in year 2000 was considerably high for the groups of low and, particularly, middle developed countries (home of 70% of the world's population). In general, we see that such contributions have tended to increase over time – except for the group of highly developed countries, where they have tended to remain at relatively low levels. Considering all world countries together as a single entity (i.e. "All countries"), we observe increases in the contribution of within-country inequality for the SHDI and its three basic sub-components from 2000 to 2015. Such increase in the contribution of the within-country inequality components has gone in tandem with declines in the levels of inequality for the SHDI and its sub-indices (see main paper's Figure 2).

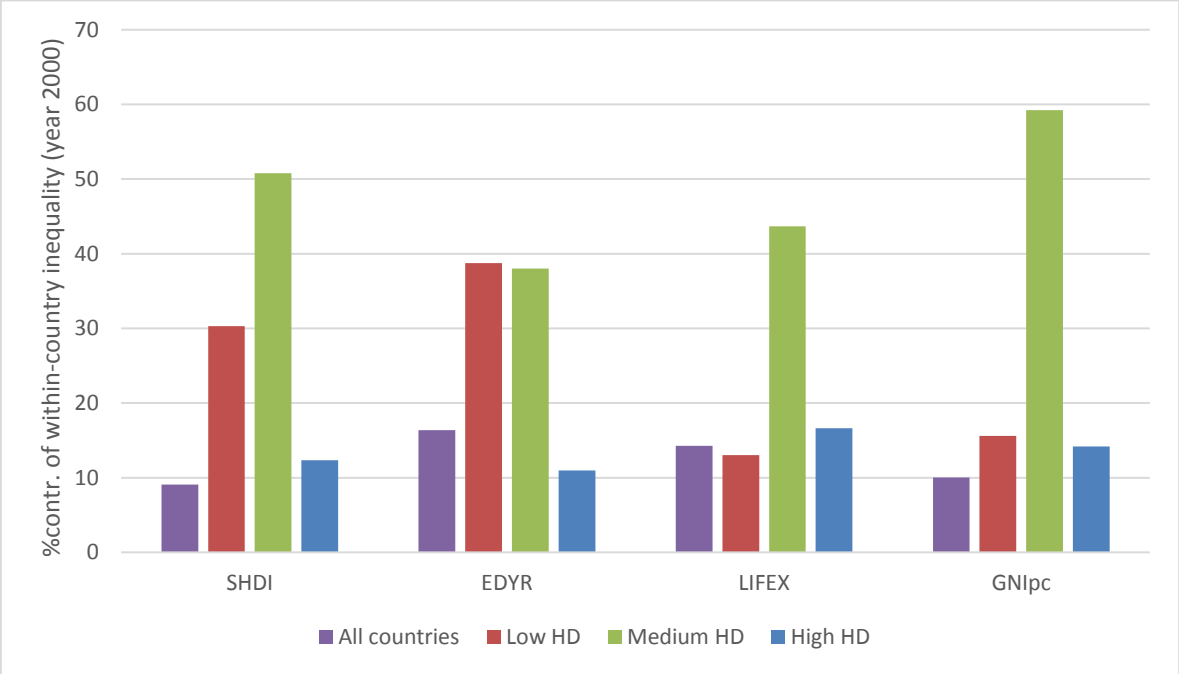

Figure S1. Percent contribution of within country inequality to total inequality for the SHDI and its three subcomponents in different groups of countries (year 2000). To define whether a country belongs to the group of Low, Medium or Highly developed countries we look at its HDI in year 2008 (the middle of the observation period). If such HDI is below (resp. above) 0.55 (resp. 0.7) the country belongs to the set of Low (resp. High) developed countries. Whenever it falls between 0.55 and 0.7, it is a Medium developed country. Source: Authors' elaboration based on information from the Global Data Lab.