


---

This is the **accepted version** of the journal article:

Permanyer, Iñaki; Suppa, Nicolai. «Racing ahead or lagging behind? Territorial cohesion in human development around the globe». *Regional Studies*, Vol. 56, Núm. 12 (2022), p. 2086–2101. DOI 10.1080/00343404.2022.2038785

---

This version is available at <https://ddd.uab.cat/record/271441>

under the terms of the  <sup>IN</sup>  
COPYRIGHT license

**Artículo de revista:**

PERMANYER, Iñaki; SUPPA, Nicolai (2022) "Racing ahead or lagging behind? Territorial cohesion in human development around the globe". *Regional Studies* (online: 04 Apr 2022): 1-16 (ISSN: 0034-3404).  
<https://doi.org/10.1080/00343404.2022.2038785>

# Racing ahead or lagging behind? Territorial cohesion in human development around the globe.

Iñaki Permanyer<sup>\*,†,‡</sup>      Nicolai Suppa<sup>†,§</sup>

November 26, 2021

## Abstract

This paper investigates whether global improvements in human development involve subnational regions in a territorially cohesive way. We use a subnational human development index for 1778 regions within 163 countries over three decades, and propose measures for relative over- and under-performance. We find that frequently observed reductions of within-country inequality are not necessarily accompanied by reductions in under- and over-performance. Moreover, from a global perspective, we detect the presence of a non-negligible set of under-developing subnational regions spanning across 30+ countries that fail to catch-up with the world average human development.

**Keywords:** human development; inequality; regional disparities; territorial cohesion

**JEL Codes:** I31, O15, D63, R11

---

\*Corresponding author.

<sup>†</sup>Centre d'Estudis Demogràfics (a member of CERCA, Generalitat de Catalunya), Carrer de Ca n'Altayó, Edifici E-2, Campus de la UAB, 08193 Cerdanyola del Vallès, Spain and EQUALITAS; contact: [iper-manyer@ced.uab.es](mailto:iper-manyer@ced.uab.es) and [nsuppa@ced.uab.es](mailto:nsuppa@ced.uab.es).

<sup>‡</sup>ICREA, Passeig Lluís Companys 23, Barcelona, Spain

<sup>§</sup>Oxford Poverty and Human Development Initiative (OPHI), Oxford Department of International Development, University of Oxford.

**Data availability:** The data underlying this article are available in the *Global Data Lab* at [https://globaldatalab.org/shdi/download\\_files/](https://globaldatalab.org/shdi/download_files/) and can be accessed under “Version 5.0”.

**Funding information:** Permanyer gratefully acknowledges funding of the European Research Council (ERC-2019-CoG-864616, HEALIN project), the Spanish Ministry of Economy and Competitiveness “Ramón y Cajal” Research Grant Program under Grant RYC-2013-14196; and the Spanish Ministry of Economy and Competitiveness National R&D&I Plan under Grant RTI2018-096730-B-I00. Suppa also gratefully acknowledges funding of the Spanish Ministry of Science, Innovation and Universities Juan de la Cierva Research Grant Programs (IJC-2017-33950).

**Conflict of interest disclosure** We have no conflict of interest to disclose.

**Acknowledgments** This paper greatly benefited from comments made by Maria Emma Santos, participants of the RUE seminar at the University of Barcelona, WEAI 2021, ECINEQ 2021, ISQOLS 2021 and two anonymous referees. All remaining errors are our own.

# 1 Introduction

Since the late 20<sup>th</sup> century, living conditions of people around the world improved considerably. In most regions of the world humans now enjoy longer lives (Riley, 2001), better education (Morrisson and Murtin, 2009) and higher living standards (Easterlin, 2000). While much is known about the average improvements of these dimensions over time, much less is known about the inequality part of the story. In particular, a spatial view on inequalities in ‘living conditions’ or ‘human development’ across geographical or administrative units within countries, has recently received considerable attention from researchers and policy-makers—and for good reasons.<sup>1</sup> First, individuals with similar socio-economic characteristics often concentrate in space, suggesting that increases in spatial inequality tend to be positively associated with greater levels of interpersonal inequality (Kanbur and Venables, 2005, Lessmann, 2014). Second, high levels of spatial inequality often come along with political and ethnic tensions (Ezcurra and Rodríguez-Pose, 2017), and could eventually lead to conflicts and civil wars (Buhaug *et al.*, 2011, Deiwiiks *et al.*, 2012, Ezcurra and Palacios, 2016). Unsurprisingly, social- and territorial-cohesion goals have been explicitly included in the agendas of major recent development endeavors, like the ‘EU 2020 Strategy’ or the sustainable development agenda, where goal #10 invites to ‘reduce inequality within and among countries’.

Attempts to assess the extent to which the living conditions across countries’ subnational units are evolving in a territorially cohesive way are hindered by the lack of both appropriate measures and the underlying data. On the one hand, currently existing methodological approaches, which include ‘convergence’ and ‘inequality’ analyses (e.g. Johnson and Pappageorgiou, 2020, Cowell, 2011), have certain shortcomings as they fail to capture some intuitive notions one might want to take into consideration when assessing countries’ territorial cohesion. While the former explore whether ‘lower developed’ regions are growing at a faster rate than the others, the latter examine the spread of the inter-regional distribution. However, none of them takes into consideration what is actually happening at the lower and the upper tails of the distribution. On their own, the existing methods are unable to ascertain whether and to what extent some specific regions are lagging behind or racing ahead of the rest. This deficiency is problematic as social progress might be easily overstated—even if improvements in both averages and inequalities are taken into account. Several authors argued forcefully that the situation of the worst-off is relevant for any evaluation of social arrangements or progress (e.g., Rawls, 1999, Sen, 1999, 2009). Indeed, in the sustainable development agenda the United Nations explicitly ‘pledge that no one will be left behind’ in the development process.<sup>2</sup> On the other hand, estimating the evolution of the living con-

---

<sup>1</sup>Throughout the paper, the expressions of ‘living conditions’, ‘dimensions of human well-being’, and ‘human development’ will be used interchangeably. In practice, these notions will be operationalized via the United Nations’ Human Development Index (HDI) — see below.

<sup>2</sup>See <https://sustainabledevelopment.un.org/post2015/transformingourworld>.

ditions across countries' subnational units is not an easy task. Data limitations have often restricted spatial inequality analyses to a handful countries, in particular when it comes to inequality in more comprehensive well-being measures (see [Grimm et al. 2008](#), [Permanyer et al. 2015](#) for examples from the global south and [Veneri and Murtin 2018](#), [Iammarino et al. 2019](#) for examples from the global north). Therefore, a truly global perspective analyzing countries' territorial cohesion has not been implemented yet. In this paper, we aim at making substantive contributions on both fronts.

On the methodological side, we propose new 'under-development' and 'over- development' measures that explicitly focus on what is happening at the bottom and the top of the inter-regional distribution, respectively. Together, these measures are meant to assess the extent to which countries' subnational regions are performing exceedingly better or worse than the national average, thus revealing whether living conditions are distributed in a territorially cohesive way. Conceptually, they are straightforward adaptations of well-known 'poverty' ([Foster et al., 1984](#)) and 'richness' ([Bose et al., 2014](#)) indices that are commonly applied to study what happens at the bottom and top tails of inter-individual income distributions. In the same way as poverty and richness measures enhance and complement the insights provided by inequality measures in the context of income distribution analysis, the new over- and under-development measures proposed in this paper are a useful complement to the 'convergence' and 'inequality approaches' commonly applied in the regional studies literature.

On the data side, we aim for a global coverage in our analysis of territorial cohesion in human development. For this purpose, we take advantage of the Subnational Human Development Index (SHDI) Database ([Smits and Permanyer, 2019](#)), which contains estimates of the United Nations' Human Development Index and its sub-components across more than 160 countries and 1600 subnational units representing more than 99% of the world's population. Since such estimates are available from 1990 onwards, we are able to cover almost three decades in our analyses. More specifically, the richness of the dataset allows us to document (i) the levels and trends of inequality in human development, (ii) the levels and trends of over- and under-development, and (iii) the contribution of over-/under-development to inequality, and that for both within countries and in the world as a whole. Put differently, we will examine whether, and to what extent, the fact that some subnational units lag behind or race ahead of the rest contributes to the inequality in human development observed both from a local (i.e. country-level) and a global perspective. Our findings suggest that frequently observed reductions of within-country inequality are not necessarily accompanied by reductions in under- and over-performance. Moreover, from a global perspective, we detect the presence of a non-negligible set of under-developing subnational regions spanning across 30+ countries that, in the last two decades, have failed to catch-up with the world average human development—a finding that remained concealed so far due to the insufficient granularity of currently existing databases and the limitations

of ‘convergence’ and ‘inequality’ techniques.

The remainder of the paper is structured as follows: section 2 provides more background information on related research, section 3 introduces the applied methods and section 4 provides more details about our data. Section 5 presents the results for our within-country analysis and section 6 for our global analysis. Finally, section 7 offers some concluding remarks.

## 2 Background

The present study relates to different lines of previous research. On the conceptual side, one of the major criticisms leveled against the HDI is its neglect of inequalities. As a consequence some studies propose modifications of the measure itself to account for inequalities (e.g., [Foster et al., 2005](#), [Seth, 2009](#)). Other studies instead seek to calculate the HDI for subgroups to allow various comparative analyses subsequently.<sup>3</sup> For instance, [Grimm et al. \(2008, 2010\)](#) calculate the HDI for income quintiles in 32 countries, [Harttgen and Klasen \(2011\)](#) for internal migrants and non-migrants in 16 low-income countries, whereas [Harttgen and Klasen \(2012\)](#) propose a method to proxy the HDI at the household-level, which is illustrated using 15 countries. Moreover, [Permanyer \(2013\)](#) proposes a municipality-level HDI, which is illustrated using census data from Mexico, whereas [Permanyer et al. \(2015\)](#) apply this approach to 13 African countries and document inequalities in human development at the national level. We follow these papers in exploiting subnational variation in human development to incorporate inequality into our analysis.

Accounting for within-country inequality along these lines, however, tends to restrict in the analysis to rather few countries, as is already evidenced by the previously referenced papers. Moreover, several recent studies, explore regional disparities in a similar fashion, however, using measures of well-being other than the HDI (e.g., [Ballas et al., 2017](#), [Veneri and Murtin, 2018](#), [Peiró-Palomino, 2018](#), [Pinar, 2018](#), [Iammarino et al., 2019](#), [Ayala et al., 2020](#)). These studies on the other hand are usually confined to industrialized countries and thus of limited geographical coverage, too. Using the SHDI database, we are able to offer an analysis of human development, which accounts for within-country inequalities and is of global scope (for a description of the data see [Smits and Permanyer 2019](#) and for some first results [Permanyer and Smits 2020](#)).

On the methodological side, the concept of convergence has been operationalized in different ways. Two particular prominent approaches,  $\beta$ -convergence and  $\sigma$ -convergence, have been explored in research on the convergence of per capita income (e.g., [Sala-i-Martin, 1996](#), [Durlauf et al., 2009](#)). According to the first approach poor countries grow faster than rich countries, whereby they are effectively catching up. Essentially, one can test for  $\beta$ -

---

<sup>3</sup>Another approach is to examine multidimensional inequality at the global level using national HDIs as pursued in [Decancq et al. \(2009\)](#).

convergence via regressing the growth rate, e.g., of income per capita on its initial levels. A significant negative coefficient would imply initially richer countries to grow at a slower pace. Instead, the so-called  $\sigma$ -convergence measures requires the standard deviation of a particular outcome variable across countries or regions to decrease over time. Since  $\beta$ -convergence is necessary but not sufficient for  $\sigma$ -convergence and some further results, many authors recommend to directly investigate the variance; for more details see [Johnson and Papageorgiou \(2020\)](#), a recent survey on the convergence of income per capita on the international level. Note, that these methods have also been applied to quality of life indicators (e.g., [Mazumdar, 2003](#), [Neumayer, 2004](#)) including the HDI ([Jordá and Sarabia, 2015](#)).

While the standard deviation reflects dispersion, inequality measures satisfy several desirable axioms, which may prove useful in the analysis of convergence (e.g., [Salas, 2002](#), [Durlauf et al., 2009](#)). Indeed, inequality measures are well-understood and several measures allow instructive decompositions, e.g., into contributions of subpopulations or within- and between components ([Cowell, 2011](#)). By now inequality measures have been applied to examine convergence in income and various other dimensions of human well-being such as education and health (see below). In this paper we propose specific over- and under-performance measures to complement the analysis of convergence in particular using inequality measures to study territorial cohesion more comprehensively. Their application can be motivated along the lines of a Rawlsian social welfare function, the SDG paradigm to leave no one behind, or to quantify the contribution of recent progress by the top-performers.

Finally, on the empirical side our paper also relates to previous research on single dimensions of human well-being such as education, health, and income. A common feature of these studies is the application of conventional inequality measures to data which provide within-country variation in one form or another to study regional disparities. Moreover, these studies aim for a high coverage of the world population to allow for truly global assessments. For instance, using a large data set of ‘macro-countries’ [Morrisson and Murtin \(2013\)](#) find for the period of 1870–2010 a hump-shaped relationship for inequalities in education as measured by Gini and Theil indices for years of education. In a similar fashion, [Jordá and Alonso \(2017\)](#) provide new mean years of schooling estimates for 142 countries for 1970–2010 and document a decreasing global inequality using Gini and Theil measures, among others.

Similarly, global inequalities in health have been analyzed extensively. For instance, [Edwards \(2011\)](#) examines inequality in length of life (i.e. age at death) for 180 countries in 1970 and 2000 and inter alia applies Gini and Theil measures. Among other things the author documents (i) a substantial decline in inequality and (ii) that around 90% of total inequality is due to within-country variation (despite an increasing importance of the between-country component). Using data of the United Nations World Population Prospects,



[Permanyer and Scholl \(2019\)](#) analyze inequality in length of life from 1950 to 2015 and document (i) a decline in inequality according to the Theil index and the variance and (ii) that most of the world variability in age-at-death can be attributed to within-country variability. Finally, research on global income inequality made recently substantial progress, see, e.g., [Milanovic \(2012\)](#) and [Lakner and Milanovic \(2015\)](#). While there is consensus on the extraordinary increases in inequality over the last two centuries, changes during the most recent decades are less clear cut and results often depend on method and data, see [Anand and Segal \(2015\)](#) for a survey.

Thus given the available evidence one may expect global inequality in subnational HDIs to decline, too. Indeed, in a first analysis of the SHDI data base [Permanyer and Smits \(2020\)](#) find the overall mean log deviation to decrease from 0.031 in 2000 to 0.0178 in 2017. In the present paper we revisit this observation of declining inequality in human development and probe whether these recent developments in human development are really a flawless story of success. More specifically, we argue that conventional inequality measures cannot reveal the full picture needed to sufficiently assess progress based on the sustainable development paradigm.

As the various lines of empirical research also reflect, accounting for within-country variation and providing a meaningful analysis of trends at the same time, tends to reduce the global coverage of countries. Consequently, the state of knowledge remains fragmentary. We complement previous research by offering a more comprehensive analysis of regional disparities in terms of a single composite measure of human well-being, accounting for both between and within variation for 163 countries over almost 30 years.

### 3 Methods

In this section we describe the basic notation that will be used throughout the paper. We start with the concepts applied in country-specific analyses first, and then proceed to the global perspective that involves comparisons across all world countries.

#### 3.1 Country level analysis

For any given country, the distribution of human development across its  $r \in \mathbb{N}_+$  subnational regions is described by a vector of achievements  $x = (x_1, \dots, x_r)$  and population shares  $p = (p_1, \dots, p_r)$ , where  $x_i \in [0, 1]$  and  $p_i \in [0, 1]$  correspond to the level of human development (as measured by the United Nations' HDI) and the population share of region  $i$ , respectively. As the human development index is frequently used as a measure of human well-being, our analyses employ population weights in all instances, including the measures of inequality, over- and under-development, and also when aggregating across countries. In this paper, we use one of the most popular inequality measures, the Gini index ( $G$ ), which is defined

as follows

$$G(x, p) = \frac{\sum_i \sum_j p_i p_j |x_i - x_j|}{2\mu} \quad (1)$$

where  $\mu = \sum_i p_i x_i$  is the national-level mean. As is well-known, the values of  $G$  are bounded between 0 and 1, which are observed in the cases of perfect equality (i.e. all regions have the same level of HDI) and extreme inequality (i.e. all regions except one have the lowest possible HDI level of 0), respectively.

The extent of *under-development* in a given country relative to the national mean is defined as

$$U^\gamma(x, p) = \sum_i p_i \max \left\{ \frac{\underline{z} - x_i}{\underline{z}}, 0 \right\}^\gamma \quad (2)$$

where,  $\underline{z} = \mu - a \geq 0$ , with  $a \in (0, \mu]$ , and  $\gamma$  is a non-negative parameter. In this way, subnational regions with an HDI level that is  $a$  units below the corresponding national HDI, will be considered as ‘under-developed’.<sup>4</sup> The set of regions within a given country whose development level falls below this threshold (i.e. the ‘under-developed regions’) will be denoted as  $\mathcal{U}$ . When  $\gamma = 0$ ,  $U^0$  is analogous to a headcount poverty measure, and measures the share of the population in that country that lives in regions with a human development level below the threshold given by  $\underline{z}$ . In general, for  $\gamma > 0$ ,  $U^\gamma$  simply measures the population weighted average of the ‘under-development gap’ ( $\max(\underline{z} - x_i)/\underline{z}, 0$ ) raised to the power of  $\gamma$  across subnational units. Thus, our class of under-development measures is formally equivalent to the Foster-Greer-Thorbecke class of poverty measures (Foster *et al.*, 1984). When  $\gamma = 0, 1$  and  $2$ ,  $U^\gamma$  are analogous to the ‘headcount ratio’, the ‘poverty gap measure’ and the ‘squared gap measure’, respectively. To better document the normative foundations of these measures, we detail in the appendix some of the axioms that this class of measures satisfies, along with some explanation why we consider these axioms important in the present context. Observe that  $U^\gamma(x, p)$  is a purely *relative* measure of under-development: it captures the extent to which some regions are lagging behind the national average, irrespective of the absolute values of the distribution (i.e. a highly developed country can have an ‘underdeveloped’ region with a certain development level that would not qualify as such in the distribution of other, less-developed, countries). In addition to this, also note that population shares are taken into account, i.e. a more populous region

---

<sup>4</sup>We note that cutoffs  $\underline{z}$  are related to the mean performance ( $\mu$ ), and may be set in different ways. The at-risk-of-poverty rate used in the European Union, for instance, relies on a *proportion* of the median income. At least two factors seem relevant in this context. First, the empirical distribution of the underlying welfare variable requires careful inspection with respect to (i) commonly observed variation in the cross-section, (ii) existence and implications of time trends (in particular for bounded variables), but also (iii) the role historically rare observations. Second, simple thresholds are easier to communicate to the public, which is important to facilitate the process of setting cutoffs in practice. Plain HDI differences are reported in units of the HDI itself and thus are meaningful, which allows illustrations using cross-country comparisons, for instance.

lagging behind increases *ceteris paribus* our measures of under-development.

Using a similar notation, we can define the extent of over-development associated to the distribution as follows

$$O^\gamma(x, p) = \sum_i p_i \max\left(\frac{x_i - \bar{z}}{\bar{z}}, 0\right)^\gamma \quad (3)$$

where  $\bar{z} = \mu + b$ , with  $b \in (0, 1 - \mu]$ .  $O^\gamma(x, p)$  should be interpreted as the extent of relative over-development we observe in a given country, where the parameter  $b$  specifies the minimum over-performance a region is required to have with respect to the national mean for being identified as ‘racing ahead’. The set of regions within a given country whose development level is above the over-development threshold  $\bar{z}$  will be denoted as  $\mathcal{O}$ . Like in the previous case, when  $\gamma = 0$ ,  $O^0$  measures the share of the population in a given country that lives in regions with a human development level above the threshold given by  $\bar{z}$ . Likewise, when  $\gamma = 1$  and  $\gamma = 2$ ,  $O^\gamma$  is a population-weighted average of ‘over-development gaps’ and ‘squared over-development gaps’, respectively. The class of over-development measures proposed here ( $O^\gamma$ ) mirrors the under-development one ( $U^\gamma$ ), and is formally equivalent to some of the ‘richness indices’ (i.e., real-valued functions indicating the extent of economic richness in a given society) presented in (Bose *et al.*, 2014).

In this paper we want to explore the relationship between over-/ under- development and the inequalities in human development we observe across countries and in the world as a whole. In particular, we aim to assess the extent to which the phenomena of over- and under-development contribute to the existing inequality levels. For that purpose, we decompose the Gini index as  $G = \sum_i G_i$ , where each  $G_i$  is defined as

$$G_i = \frac{\sum_j p_i p_j |x_i - x_j|}{2\mu}. \quad (4)$$

Observe that  $G_i$  can be either interpreted as the ‘degree of diversity’ or region  $i$  from all other regions in the country (see Ceriani and Verme (2015), Kendall and Stuart (1958)), or as the contribution of that region to the extent of HDI inequality in that country. This decomposition has been chosen for its simplicity and normatively desirable properties<sup>5</sup> (see Ceriani and Verme (2015)). Taking advantage of this decomposition, we define the contribution of under- and over-development to countries’ HDI inequality as

$$C_{\mathcal{U}} = \frac{\sum_{i \in \mathcal{U}} G_i}{G} \quad (5)$$

$$C_{\mathcal{O}} = \frac{\sum_{i \in \mathcal{O}} G_i}{G} \quad (6)$$

---

<sup>5</sup>These properties are additivity, continuity, anonymity, symmetry, translation invariance, linear homogeneity.

### 3.2 Global level analysis

In the empirical section of the paper we also explore the global distribution of human development across and within world countries. The Gini index of that distribution can be written as

$$\Gamma = \frac{\sum_c \sum_d \sum_i \sum_j p_{ci} p_{dj} |x_{ci} - x_{dj}|}{2M} \quad (7)$$

where  $p_{ci}$  is the population share of region  $i$  in country  $c$  (with respect to the world population),  $x_{ci}$  is the corresponding level of human development as measured by the SHDI, and  $M$  is the world mean of the SHDI distribution.<sup>6</sup>

To define the measures of under- and over-performance associated to the global SHDI distribution we need to define global under- and over-development thresholds as fractions of the world SHDI average performance ( $M$ ). Such thresholds determine what are the sets of global under- and over-developed subnational regions (i.e. akin to the  $\mathcal{U}, \mathcal{O}$  sets defined at the country level). After that, one simply needs to apply equations (2) and (3) to the global SHDI distribution. The key difference between global or country-level measures of under- and over-performance is the mean with respect to which the relative performance is assessed. Having defined the global over- and under-development thresholds, we can apply equations analogous to (5) and (6) to the global SHDI distribution to assess how much over- and under-performing regions contribute to the global SHDI inequality  $\Gamma$ .

## 4 Data

This paper uses data of the Subnational Human Development Index Database version 5.0, which is freely available online, see [Smits and Permanyer \(2019\)](#). Methodologically, the SHDI is a translation of UNDP's official HDI 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,

---

<sup>6</sup>It is worth pointing out that other well-known classes of inequality measures like the Theil index or the Mean Log Deviation (which are commonly used to assess global inequality and its decomposition across and within world countries) are not fit for purpose here. The main aim of the paper is not to decompose global SHDI inequality in its within- and between-country components, but to assess how much do over- and under-development contribute to inequality (both within countries and across the globe). The above-mentioned indices do not lend themselves to the decompositions shown in equations (4), (5) and (6).

([www.globaldatalab.org/areadata](http://www.globaldatalab.org/areadata)) and (iii) the HDI database of the United Nations Development Program (UNDP, <https://hdr.undp.org/data>).

In many low-income countries, the subnational values of life expectancy and the GNI per capita were not available, so they were estimated on the basis of related indicators. For the GNI, this was done on the basis of data for the International Wealth Index (IWI), which measures the wealth level of households on the basis of the household's possession of consumer durables, access to basic services and housing quality in a cross-nationally and cross-temporally comparable way (Smits and Steendijk, 2015). For life expectancy, estimation was based on under 5 mortality (U5M) data. In both cases, regression models were used, that were based on data at the national level (the adjusted  $R^2$  of these regression models were 82.6 and 89.1 respectively).

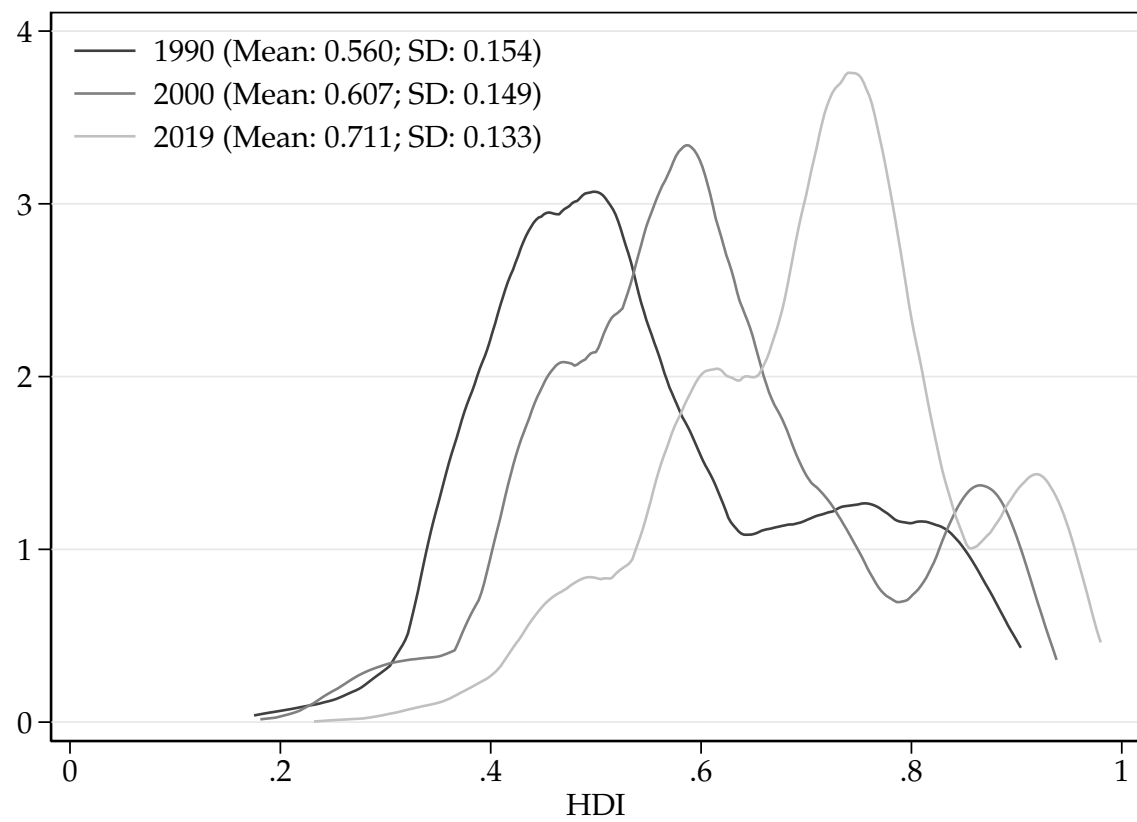
The use of indicators derived from household surveys for low- and middle-income countries means that for these countries only data is available for the years in which surveys were held. Subnational indicator values for other years therefore are estimated on the basis of interpolation and extrapolation from the survey years. Data validation analyses of Smits and Permanyer (2019) indicate that the errors due to using interpolated and extrapolated data are small.

Before entering into the computation of the SHDI, the subnational indicators are scaled in such a way that their (population weighted) national averages for a given year coincide with the national UNDP values for that year. In this way, the constructed indicators and the SHDI index are at the national level in line with their official UNDP values.

The computation of the Subnational Human Development Index, first requires to estimate the education, health and standard of living subcomponents ( $e_i, h_i, s_i$ ) and scale them between 0 and 1. This is done through the standard normalization used by UNDP in the construction of the HDI – whereby a given indicator  $X$  is transformed into a  $[0,1]$  scale via the transformation  $X^* = (X - X_{min}) / (X_{max} - X_{min})$ , with  $X_{min}$  and  $X_{max}$  being the lower and upper goalposts indicating the “natural zeros” and “aspirational targets,” respectively, from which component indicators are standardized. The values of those goalposts do not change over time and coincide with those used in the construction of the official HDI. Finally, mimicking the most recent definition of UNDP's HDI, the Subnational Human Development Index for each subnational area ‘ $i$ ’ is defined as  $SHDI_i^m = \sqrt[3]{h_i e_i s_i}$ . 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).

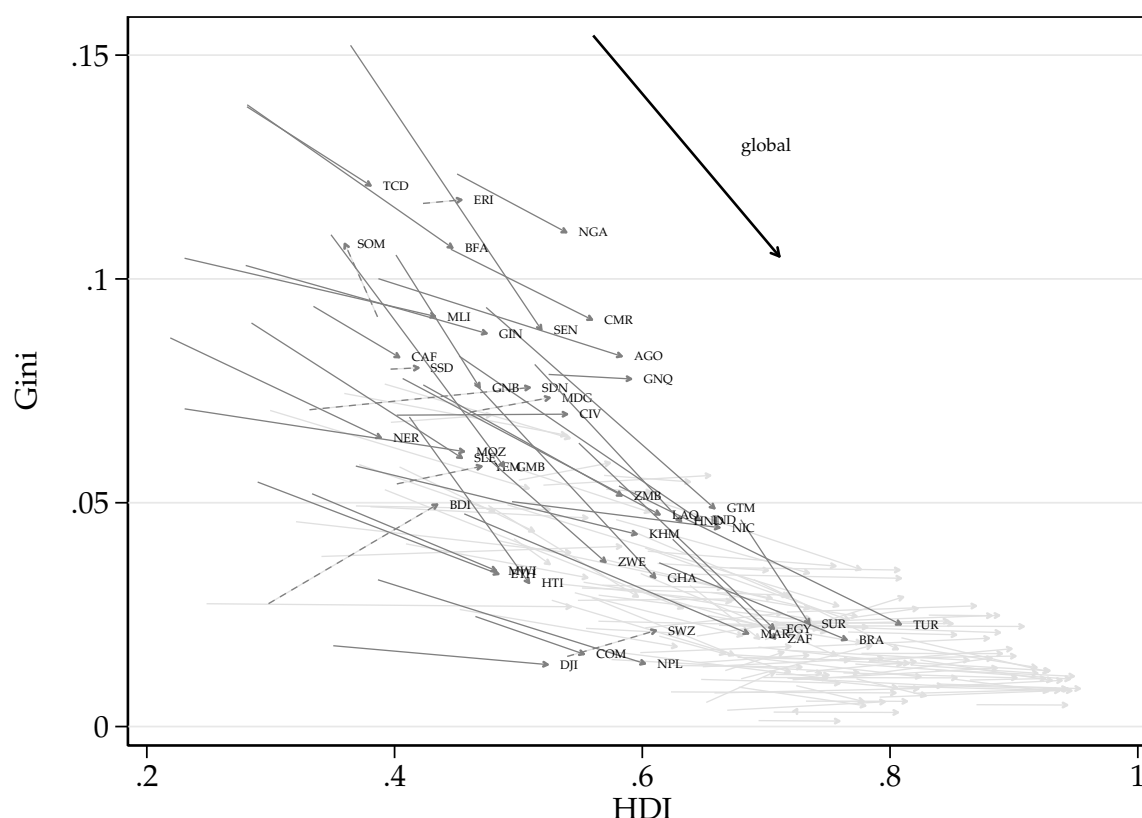
For some countries our data base contains entries since 1990, whereas other countries only join later. Data is available for most countries since 2000. For more information on the available years see table A.1. Nonetheless, the majority of our analyses rely on the entire period of observation to shed some light on long-run developments as well. Occasions where varying data availability matters are appropriately flagged.

Figure 1: The distribution of subnational HDIs over time



Notes: Underlying data is an unbalanced panel. In 2000 data for eleven low-HDI countries becomes available for the first time, thus values before and after 2000 are not directly comparable. Moreover, a few further countries are added in other years, see table [A.1](#) for details

Figure 2: Joint trends in countries' HDI and SHDI inequality



Notes: Arrows show first and last observation for countries (gray arrows) and global aggregate (black arrow). Gini for countries refers to within-country inequality in human development; global Gini refers to overall inequality (i.e., inequality in human development across subnational regions all over the world). Selected countries are labeled (using alpha-3 coding) and highlighted, with decreasing inequality (solid dark gray arrows) and increasing inequality (dashed dark gray arrows) being distinguished.

It is well-known that for most countries national HDIs increase over time.<sup>7</sup> Setting the stage for our subsequent analysis, figure 1 shows the distribution of subnational HDIs for three selected years. Figure 1 reveals that not only the HDIs increased on average, but in fact the entire distribution of subnational HDIs shifts to the right over the past 29 years. Additionally, there is also some evidence in support of global convergence as the standard deviation slightly declined, too. Note, however, that the degree of convergence as measured by this standard deviation is presumably underestimating true developments as several rather poor countries become available in our data base in 2000. Indeed, these countries account for the small hump in the lower tail of the density in 2000. Yet, the standard deviation is decreasing. Nonetheless some countries experience stagnation or even decreases in their human development levels at some point.

How did inequality in human development change over this period of observation? Figure 2 shows simplified country paths in terms of both national HDI level and within-country inequality in SHDI, i.e., according to the first and last period countries are observed in.

<sup>7</sup>This data can be explored and downloaded under <http://hdr.undp.org/en/data>.

For most countries the level of the HDI tends to increase whereas SHDI inequality within countries tends to decrease over time. Moreover, global SHDI inequality (i.e.  $\Gamma$ , as defined in equation (7)) declines over the period of observation as well (black solid arrow). Yet, for 25% of our countries we observe within country inequality to increase. Most of these countries only experience minor increases at relatively low levels of inequality and high levels of HDI, though. For some countries, however, we find increasing inequality also at lower HDI levels, in particular Burundi and Somalia, but also, e.g., Eswatini, Yemen or Madagascar. Another interesting observation is that several countries achieve relatively strong reductions in HDI inequality among regions with rather modest improvements in the national HDI level (e.g., Haiti, Gambia, or Guinea Bissau) whereas other countries rather improved the overall HDI level accompanied by a more modest inequality reduction (e.g., Mozambique, Cambodia or Nicaragua).

In our subsequent analysis we re-examine the success story of this frequently observed decreasing inequality in human development and use our new over- and under-performance measures to shed more light upon the tails of the underlying distributions to offer a fuller account of territorial cohesion in human development.

## 5 Cohesion within countries

This paper explores territorial cohesion in human development from two perspectives. This section has an exclusive focus on regional disparities *within* countries, whereas section 6 adopts a truly *global* perspective, with all subnational regions effectively being pooled. The present section first details the exact specification of our over- and under-performance measures in this context and then presents the related empirical findings.

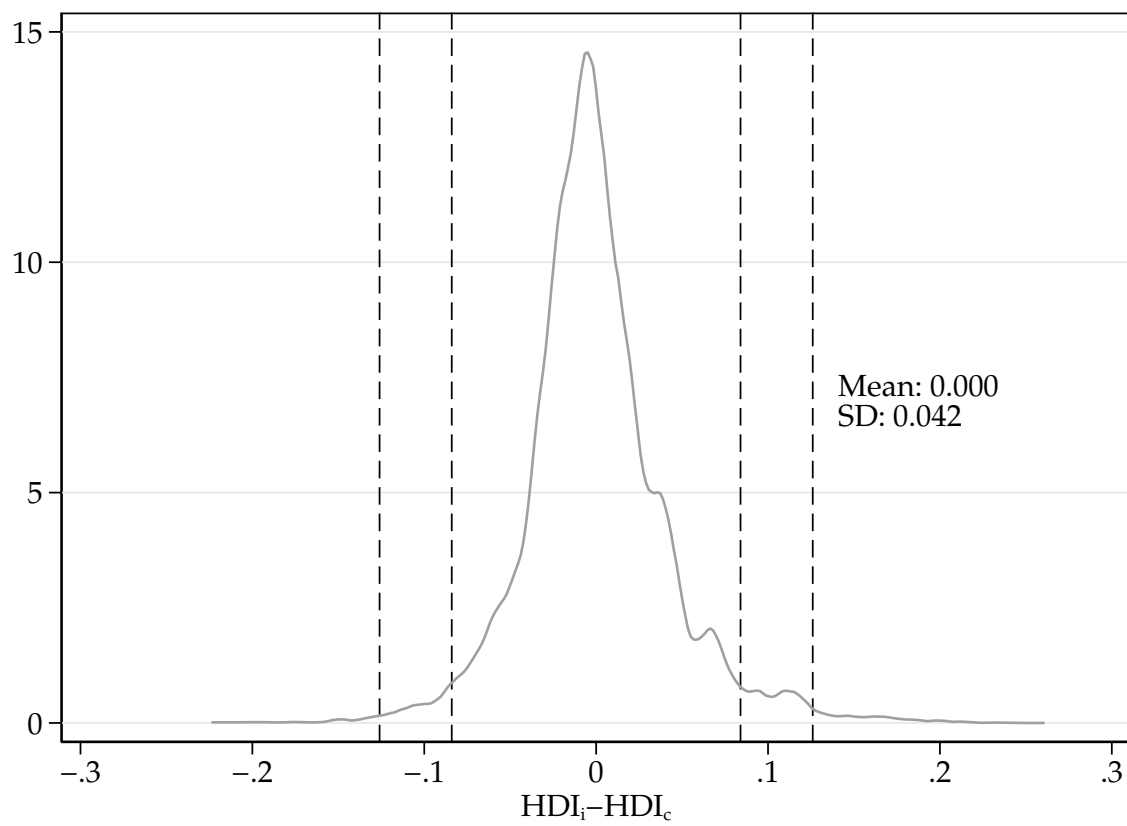
### 5.1 Specifying the relative performance measures.

**Distribution of relative performance.** In this section we are especially concerned with regions' either over- or under-performance compared with the *national* mean. Figure 3 shows the kernel density for the difference between subnational and national HDI performance (as described by  $x_i - \mu$ ) for our entire data set (all countries, all years). First, we observe a rather symmetric and relatively compressed distribution with relatively thin tails. As figure 3 shows pooled data for all available countries and all years, values below, say,  $-0.12$  and above  $0.12$  (which corresponds approximately to three standard deviations) can be considered as historically rare.

**Choosing the cutoffs.** A sense of the underlying distributions of relative performance is helpful for specifying our  $O^\gamma$  and  $U^\gamma$  measures as their cutoffs  $a$  and  $b$  both refer the relative performance. Specifically, regions potentially racing ahead or lagging behind can be found



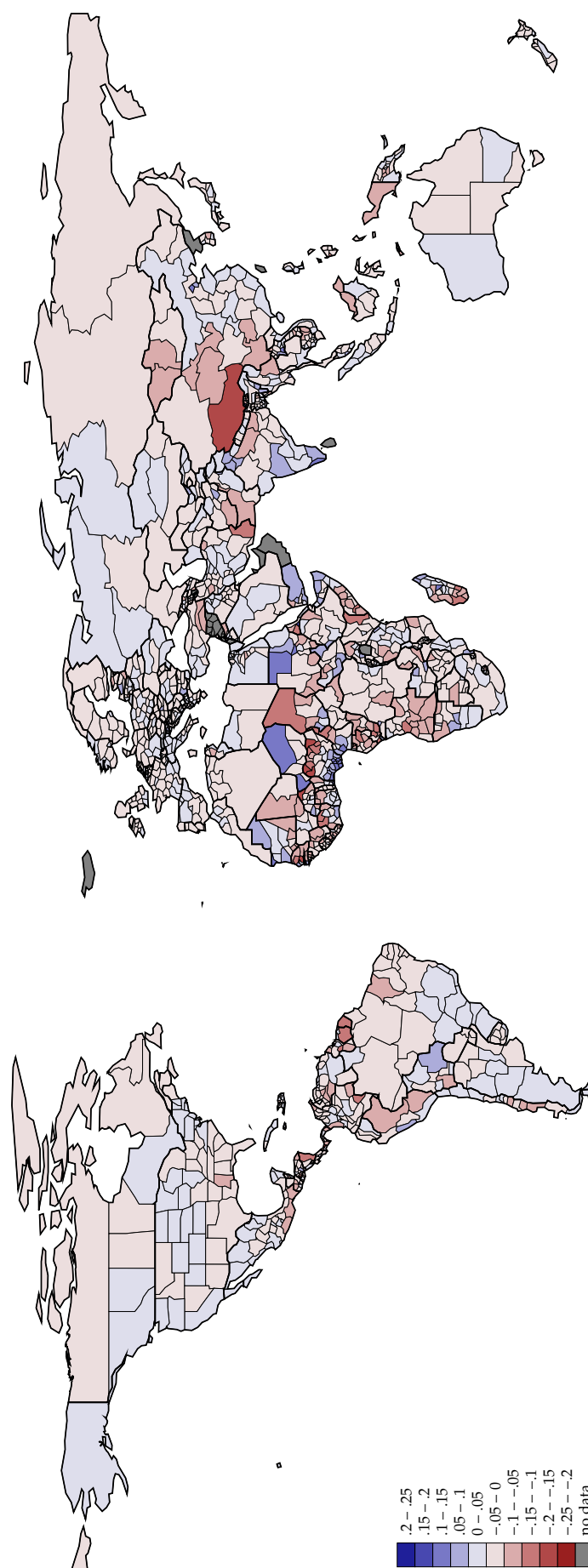
Figure 3: Kernel density for difference between national and subnational HDI



Notes: all country-year observation pooled, Epanechnikov kernel density using within country population weights, dashed lines at mean  $\pm 2$  and mean  $\pm 3$ SD, respectively.

within the tails of this distribution. We emphasize that the choice of  $a, b$  is a normative decision (similar to poverty cutoffs) and can among other things also be motivated by political priorities. For the empirical exercise in this section, figure 3 offers guidance on the implications of specific parameter choices. For simplicity, we confine our analyses to symmetric choices, even though in some applications other choices might be clearly preferable (e.g., due to the country-specific context, political priorities, etc.). Moreover, we note that there is a trade-off in choosing the parameters that has to be dealt with taking into account the specific context: While more conservative cutoffs result in lower values of the relative performance measures, lower contributions to inequality, and seemingly less relevance of over- and under-performance in general, they do allow much better to reflect and document the more extreme cases of racing ahead or falling behind. In contrast, too permissive cutoffs (which are too close to the mean performance), run the risk of making measures uninformative. For our within country analysis, we choose  $a = 0.1$  for our  $U^\gamma$  measures and  $b = 0.1$  for our  $O^\gamma$  measures as our preferred parameters. Additionally, we also report results for more permissive ( $a = 0.08, b = 0.08$ ) and more conservative ( $a = 0.12, b = 0.12$ ) thresholds, which approximately correspond to plus minus 2 and 3 standard deviations, respectively. The main findings do not however depend critically on this choice.

Figure 4: Difference between national and subnational HDI



Notes: Year of data is 2018, source of shape file is Global Data Lab.

**Geographic distribution.** The world map in figure 4 shows the subnational performance relative to the national mean (i.e.,  $x_i - \mu$ ) in 2018, and thus provides a complementary snapshot of its geographic distribution. By definition, variation around the national mean performance is observed in every country. More substantial deviations of regions from the national mean, which are therefore potentially racing ahead or lagging behind are, however, only found in 25–35 countries. Moreover, several countries have in fact both regions racing ahead and lagging behind (e.g., Nigeria, Chad or Madagascar). Furthermore, we note that most, but not all countries with strongly over- or under-performing regions are located in Sub-Saharan Africa. Finally, we can also clearly observe a gradient from coastal areas to inland in countries around the Gulf of Guinea.

## 5.2 Empirical findings

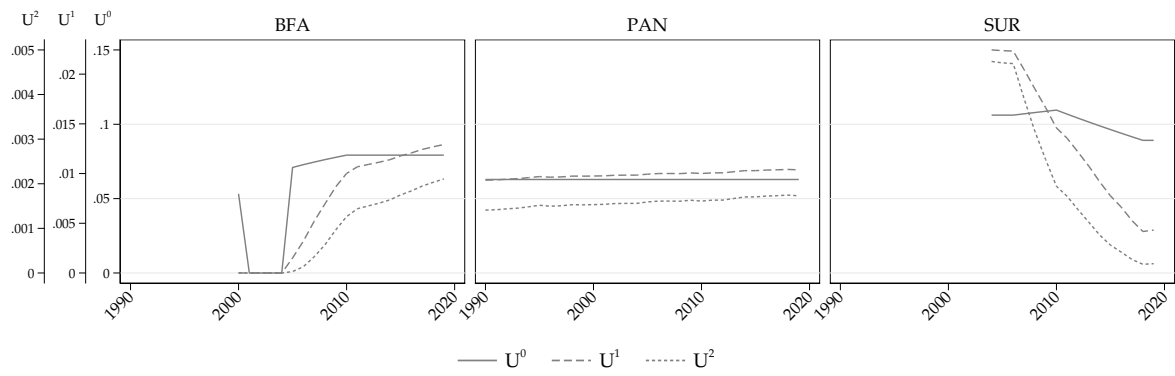
**Over- and under-performance measures.** How do our relative over- and underdevelopment measures perform in practice? Figure 5 shows both over- and under-development measures for selected countries and that for several choices of  $\gamma$  and our preferred thresholds ( $a = 0.1$ ,  $b = 0.1$ ). Results for all countries which are observed to have at least one under-performing region at some point during the period of observation, are provided by figure A.1 in the appendix.

The upper panel reveals, for instance,  $U^0$  to decline over time in some countries (e.g., Suriname or Ghana) which means that in these countries smaller proportions of their population are living in regions which are substantially under-performing in human development relative to the national mean performance. An important observation is, moreover, that the average gap in relative under-development,  $U^1$ , suggests more gradual changes than the  $U^0$  measures (see, e.g., Burkina Faso or Suriname) and is, therefore, particularly suited for monitoring progress. In Burkina Faso or Cameroon, for instance, the gap of the left behind regions to the national mean is increasing, whereas the proportion of people living in those regions remains unchanged. Together, both measures may provide useful information for national policymakers. In terms of further empirical findings, note that nowadays in several countries still have significant shares of their population living in relative under-developed regions (e.g., Cameroon and Nigeria). More generally, we observe several countries with decreases in at least one of the under-performance measures (e.g., Ghana, Cameroon or Suriname), but we also find some countries with little or no progress (e.g., Nigeria, Côte d’Ivoire or Panama) and even countries with steady or recent increases (e.g., Angola, Burkina Faso or Chad).

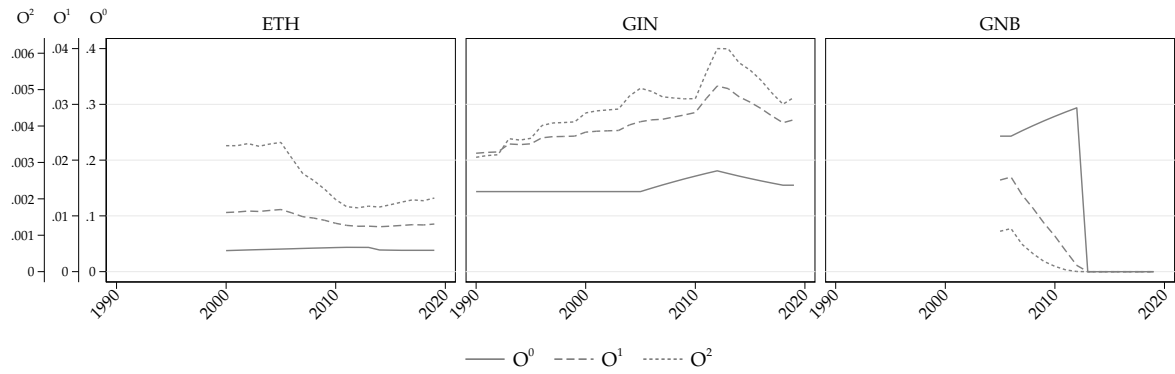
Relative over-development measures also decline over time for several countries (e.g., Guinea Bissau or Togo), only partially for some (e.g., Ethiopia) and not at all for others (e.g., Guinea or Angola). We find several countries for each of these cases according to our preferred parameter ( $b = 0.1$ ); see figure A.1 for all countries which are observed to

Figure 5: Over- and under-development in HDI (selected countries)

(a) Under-performance



(b) Over-performance



Notes: Depicted countries are found to have at least one under-developed region ( $a = 0.1$ ) or one over-developed region ( $b = 0.1$ ) during the period of observation. Multiple vertical axis refer from inner to outer axis to  $U^0, U^1, U^2$  and  $O^0, O^1, O^2$ , respectively.



with a significant level of inequality without experiencing any under- or over-performance at all, demonstrating that inequality is not necessarily accompanied by regions racing ahead or lagging behind. In sum, we observe that in many cases inequality reduction in human development are accompanied by reductions in over- or under-performance measures, but not necessarily so.<sup>9</sup> This finding suggests that one cannot infer from decreasing inequality that under-performing regions are catching up, thereby implying that a complementary analysis is required. We conclude that both measures are related in the sense that over- or under-performance require inequality, and changes in  $O^\gamma$  or  $U^\gamma$  would, *ceteris paribus*, imply changes in inequality, but not vice versa.

An alternative way to explore the nexus between both measures relies on the Gini-index decomposition outlined in section 3.1 to assess the importance of regions racing ahead or lagging behind for subnational inequality. Figure 7 (a) shows the contributions of over- and under-performing regions to the inequality in human development for all countries in 2019 for our preferred parameters (for the alternative parameters see figure A.4). The results first suggest that under- and over-performance, if present, may account for up to 30% and 40% of the observed inequality within countries, respectively.<sup>10</sup> For many countries in 2019 we observe contributions of over-performing regions between 20–30%, whereas under-performing regions in many countries contribute some 20% to within country inequality. Additionally, figure 7 also shows that according to our preferred cutoff several countries exhibit substantial contributions to inequality of both over- and under-performing regions (e.g., Nigeria, Cameroon, Angola or Burkina Faso). Finally, the upper graph in figure 7 also indicates that the lion share of subnational inequality is not driven by over- or under-performing regions.

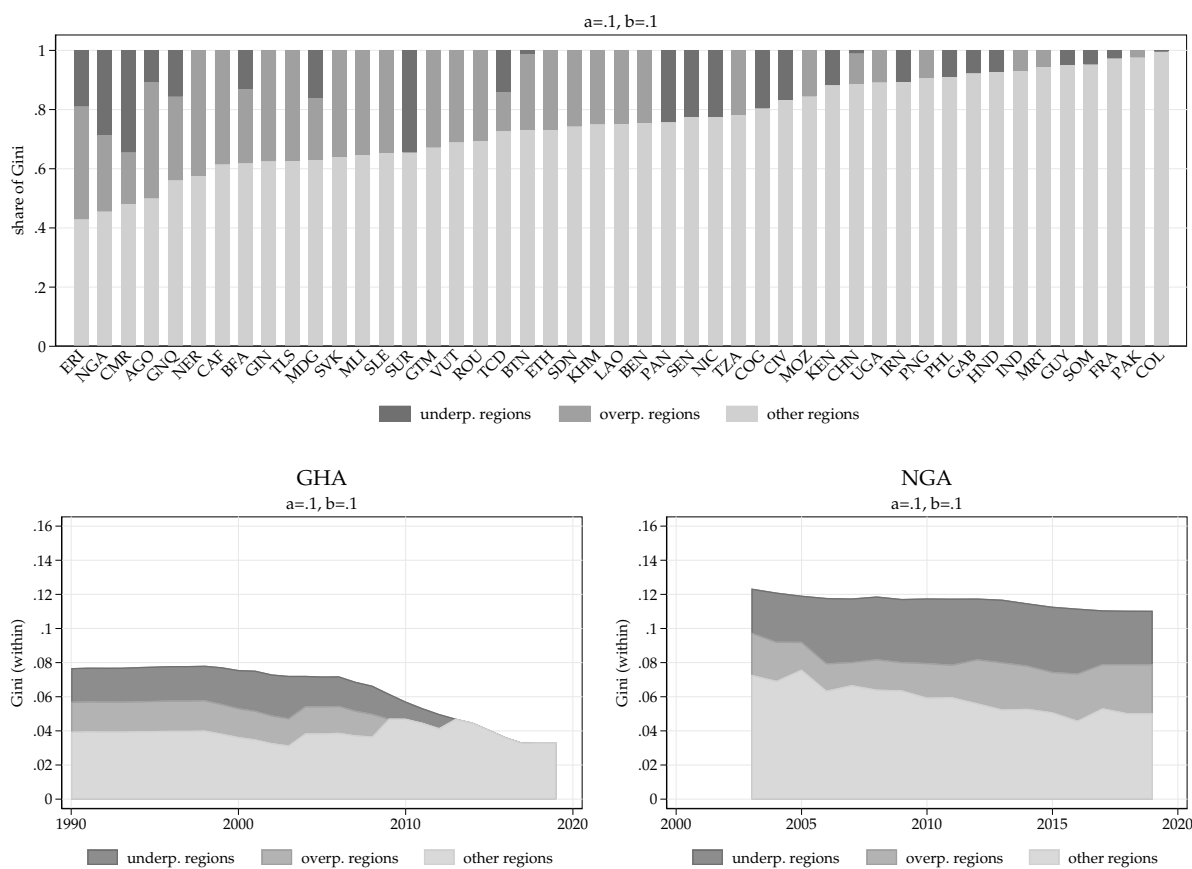
Turning to the evolution of these contributions to inequality over time, the two lower graphs of figure 7 showcase trends for two countries: Ghana and Nigeria (further examples can be found in figure A.5). For both countries we observe declining inequality of the entire period of observation, where Ghana reduces inequality in particular since the late 2000s. One difference between both countries is, however, that for Ghana the contribution of over-performing regions to inequality vanishes shortly before 2010 and that of under-performing regions shortly afterwards. In contrast, in Nigeria inequality reduction is largely associated with a decreasing contribution of regular regions, whereas contributions of over- and under-performing remain the same or even increase.

In summary, empirical findings of this section suggest that, a considerable number of countries with declining inequality exhibit regions still lagging behind or racing ahead. If present, over- and under-performing regions can easily account for 20–40% of the observed

<sup>9</sup>This finding also holds for alternative cutoffs and alternative values of  $\gamma$ . Related results are available upon request.

<sup>10</sup>Note that population size of a region matters: a less populous under-performer would have less impact *ceteris paribus*. Moreover, also note that this decomposition relies only the status of region—gaps or squared gaps do not enter the decomposition.

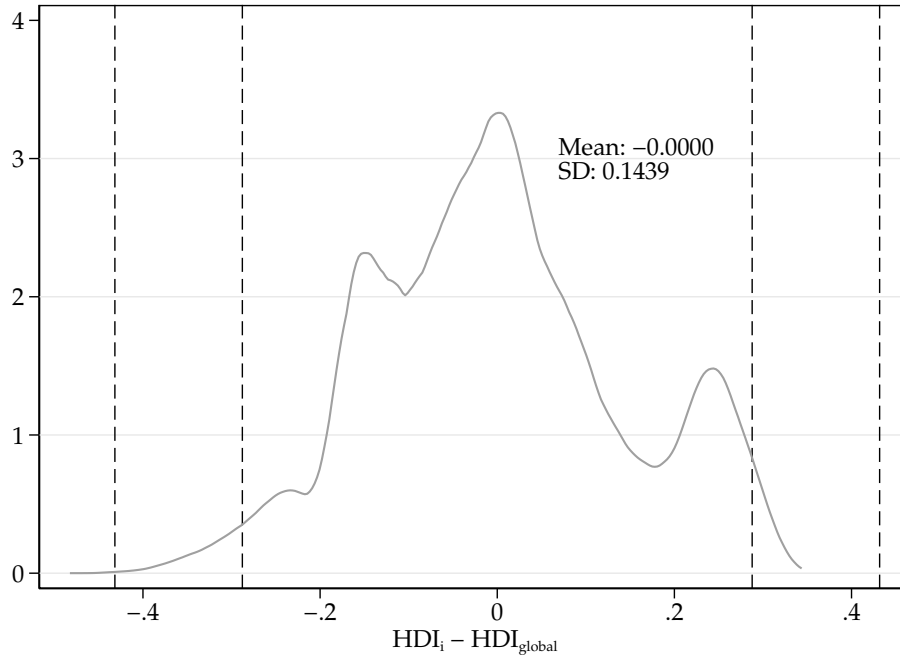
Figure 7: Regional contributions to within-country inequality (Gini)



Notes: Upper panel shows contributions of over- and under-performing regions to within country inequality in 2019; lower panels show these contributions to inequality over time for two selected countries.



Figure 8: Kernel density for difference between global and subnational HDI



Notes: all region-year observation pooled, Epanechnikov kernel density using region-to-global population weights, dashed lines at mean  $\pm$  2SD and mean  $\pm$  3SD, respectively.

inequality levels. We also find considerable heterogeneity in how exactly these trends manifest at the country-level, including polarization, stagnation, and setbacks.

## 6 Global cohesion

In this section we first specify our measures of over- under-performance relative to the global mean and then explore cohesion of the performance of subnational regions from a global perspective. This means we consider a single distribution of population weighted subnational HDIs. Inequality studied in this context, therefore, includes both within-country and between country inequality, whereas over- and under-performance measures now rely on the global, population weighted HDI.

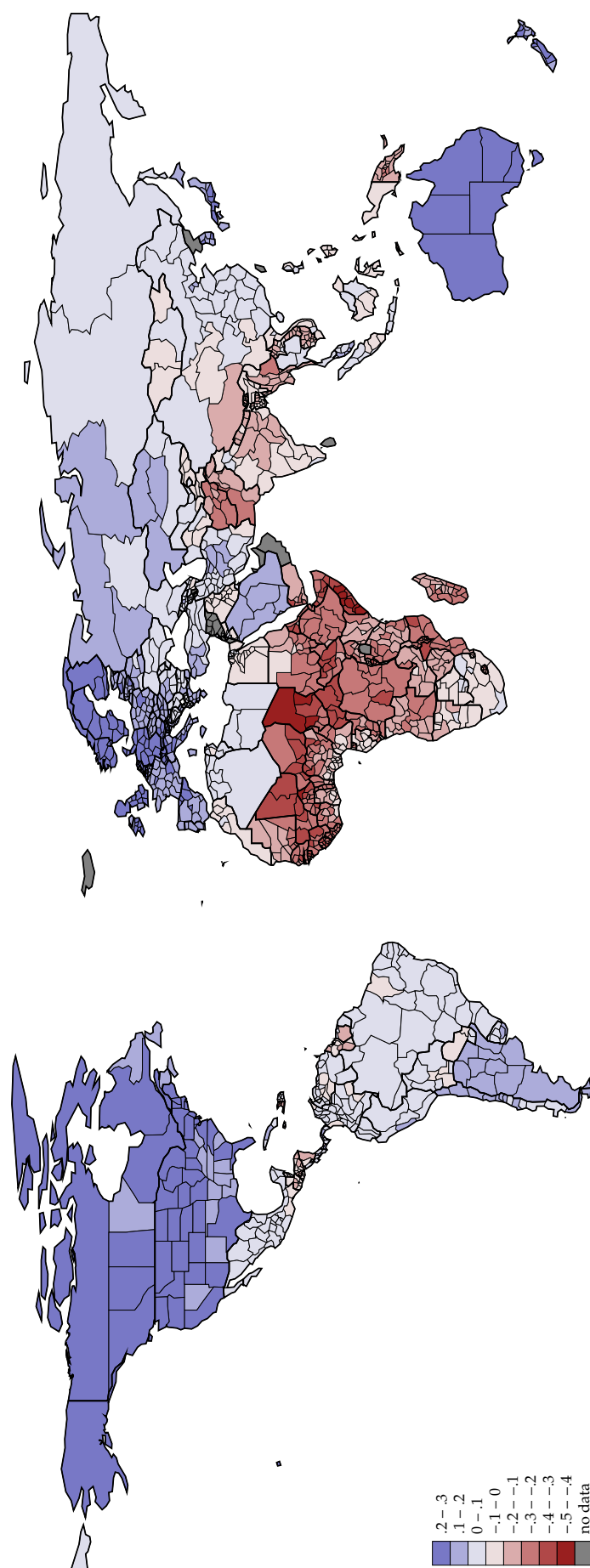
### 6.1 Specifying the over- and under-performance measures

**Distribution of relative performance.** Analogously to the previous section, we first inspect the underlying distribution of the difference between regional and global HDI to inform the choice of the thresholds of our over- and under-performance measures. Figure 8 shows the population-weighted kernel density for all regions in all years. This distribution turns out to be (i) less symmetric and in particular there is a hump on the right-hand side (reflecting many subnational regions of OECD countries), (ii) less compressed (it has a larger standard deviation), and (iii) it has shorter but somewhat thicker tails. More ex-

treme values are to be found in the tails, going beyond the mean plus minus two standard deviation approximately. These more extreme values can again be considered historically rare.

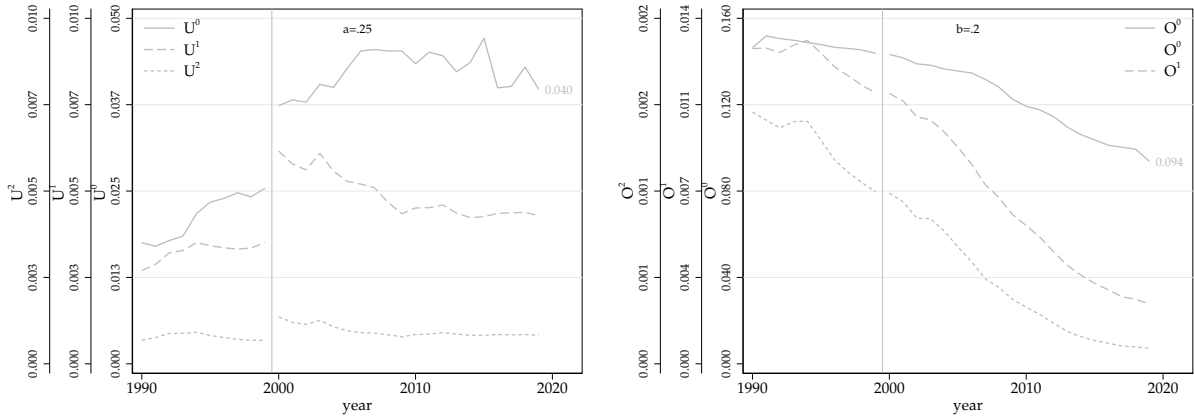
**Choosing the cutoffs.** Since our analysis of global cohesion relies on an entirely different distribution, as shown by figure 8, we also adopt different thresholds for our over- and under-performance measures to account for this different nature of the exercise. Given the slight asymmetry in the tails we also adopt slightly different thresholds for over- and under-performance measures. Specifically, for the under-performance measures our preferred cutoff is  $a = 0.25$  which again is located between 2 and 3 times the standard deviation below the mean, whereas our preferred cutoff for the over-performance measures is  $b = 0.2$ . As in the previous analyses, we also provide results for alternative cutoffs in both directions; for  $a = 0.2$  and  $a = 0.3$  and for  $b = 0.15$  and  $b = 0.25$ . Naturally, more permissive cutoffs would, e.g., induce higher contributions for inequality.

Figure 9: Difference between global and subnational HDI



Notes: Year of data is 2018, source of shape file is Global Data Lab.

Figure 10: Global over- and under-performance



Notes: Underlying data is an unbalanced panel. In 2000 data for eleven low-HDI countries becomes available for the first time, thus values before and after 2000 are not directly comparable. Moreover, a few further countries are added in other years, see table A.1 for details.

**Geographic distribution.** Turning to the geographic distribution of the subnational HDI performance relative to the global mean, figure 9 provides such a snapshot for 2018. First, we observe the well-known north-south divide and most globally under-performing regions are found in Sub-Saharan Africa, but also in South Asia and South-East Asia, and Latin America and the Caribbean. However, according to our preferred cutoffs ( $a = 0.25$ ,  $b = 0.2$ ) we find subnational regions lagging globally behind mostly in Sub-Saharan Africa with some exceptions like Afghanistan, Haiti, and Cambodia. Importantly, figure 9 also reveals that, usually, it is not entire countries that lag behind, but rather specific subnational regions within certain countries. The dividing line thus runs right through the middle of countries. More details are shown in table A.3, where we present the number of globally under-developed subnational regions within world countries, together with their corresponding population shares. In 2018 we find globally under-performing regions in human development to be scattered over some 30 countries.

## 6.2 Empirical Findings

Turning to regional over- and under-performance in human development offers a more comprehensive assessment of territorial cohesion than relying on inequality alone. Figure 10 shows how our  $U^\gamma$  and  $O^\gamma$  measures for our preferred thresholds ( $a = 0.25$  and  $b = 0.2$ ) evolve over time. The vertical line cautions to directly compare the values of our measures before and after 2000, as in this year eleven countries are observed for the first time (see table A.1 for details). For all of our over-performance measures  $O^\gamma$ , we observe a decline over the entire period (see right panel in Figure 10). This finding also holds for the alternative cutoffs, as figure A.6 in the appendix shows: in case of the stricter cutoff ( $b = 0.25$ ) we observe over-performance according to  $O^0$  to decline sharply until 2015; whereas the complementary measures  $O^1$  and  $O^2$  even suggest the remaining gap nowadays to be negli-

gible by historical standards. The more permissive cutoff ( $b = 0.15$ ) also suggests a decline over entire period, though somewhat less steep in terms of  $O^0$ . Taken together, these finding suggests that over-performance in SHDI relative to the global mean becomes both less common and less pronounced.<sup>11</sup>

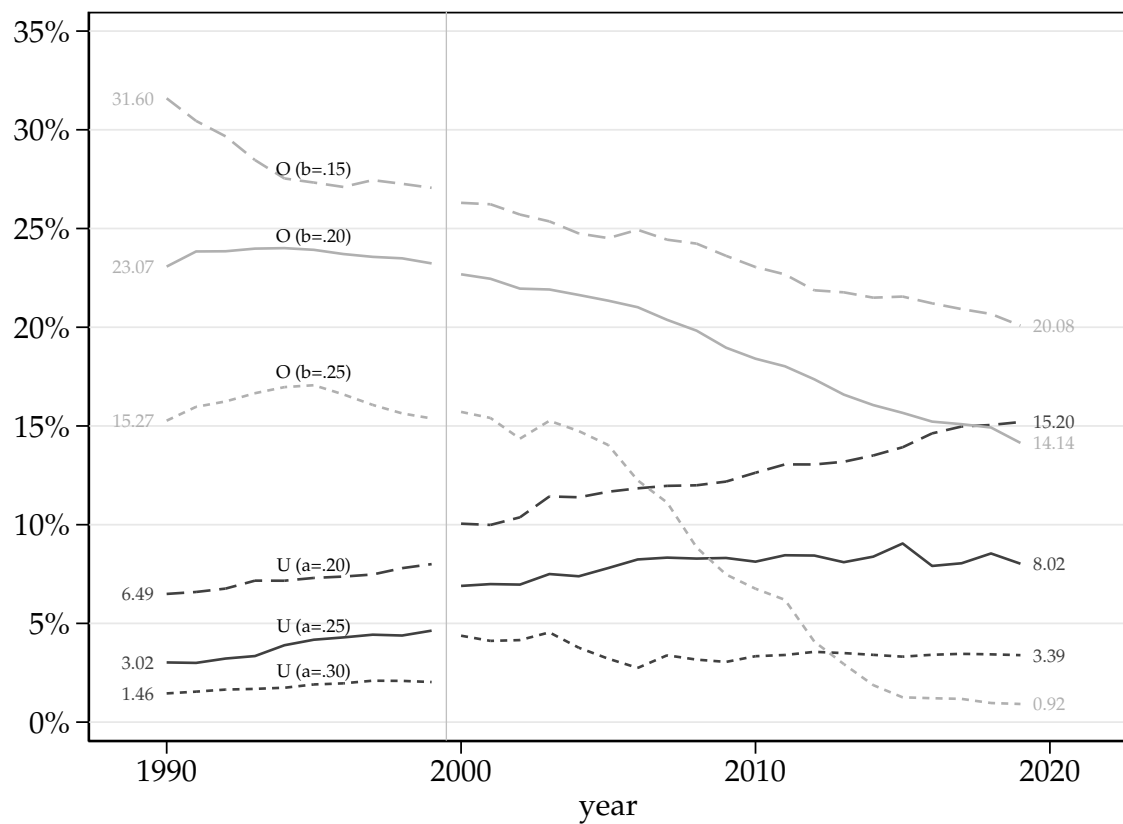
For under-performance we find little evidence in support of substantial improvements according to any of our  $U^\gamma$  measures during the period of observation (although there is some decline in  $U^1$  during the early 2000s – see left panel in Figure 10). Figure A.6 in the appendix suggests that these general trends hold for alternative cutoffs as well, even though for the stricter cutoff ( $a = 0.3$ ) we detect some period of improvement during the early 2000s, but not afterwards (for both  $U^0$  and  $U^1$ ). In contrast, for the more permissive cutoff ( $a = .2$ ) we even observe an increase according to  $U^0$  (but not for  $U^1$  and  $U^2$ ). All in all, we find little evidence in support of substantial improvement in global under-performance in terms of the SHDI, which implies that not only a considerable share of the world population is still living in globally under-performing regions, but also that these regions fail to close the gap (and the squared gap) and thus are not catching-up with the global average.

Since our global under-performance measures are defined with respect to the world SHDI mean that increases over time, the existence of a ‘long tail’ of under-developed regions does not imply that their human development levels decline over time. Rather, our findings suggest that these subnational regions are not catching up sufficiently fast, i.e. their human development levels increase equal to or below world-average speed, so the lower tail of the distribution remains ‘too long’. Since our  $U^0$  measure reports population shares and the world population covered by our data approximately amounts to 7.7 billion in 2019, we can directly infer that the under-performing regions are home to circa 308 million people according to our preferred cutoff, which is a non-negligible amount.

What is the contribution of over- and under-performing regions to global SHDI inequality? First, recall that global inequality is declining over the entire period of observation, as earlier shown by figure 2. The inequality decompositions detailed in section 3.2 allows us now to explore the role of over- and under-performing regions in this development. Figure 11 shows these contributions to inequality for our preferred thresholds (solid lines) and both alternative cutoffs (dashed lines) and largely resembles the previous figures on the measures themselves. Specifically, we observe the contribution of over-performing regions, to decline over time, irrespective the underlying cutoff. In contrast, contributions of regions lagging behind to overall inequality appear to be surprisingly stable over time. Specifically, for our preferred thresholds we find contributions of over- and under-performing regions to overall inequality to account for about 23% and 3% in 1990 and for about 15% and 8%

<sup>11</sup>While the occurrence of over-performance becomes increasingly difficult as the global SHDI approaches its natural upper bound (equal to 1), it is important to highlight that (i) it is technically feasible (i.e., there are no instances where the global over-performance threshold  $\bar{z}$  goes above 1), and (ii) as can be seen in Figure 10, the share of world population living in over-performing regions in 2019 is around 9% – a relatively large value.

Figure 11: Contributions of over- and under-performance to global inequality



Notes: Underlying data is an unbalanced panel. In 2000 data for eleven low-HDI countries becomes available for the first time, thus values before and after 2000 are not directly comparable. Moreover, a few further countries are added in other years, see table A.1 for details.

in 2019. We therefore conclude that the composition of global total inequality has shifted over time.

In summary, results in this section show that global over-performing regions in human development previously played an important role in ‘making for global inequality’, but their importance is declining over time. Additionally, our under-performance measures reveal a considerable population to live in regions which fail to close the gap with global average performance and are, therefore, still lagging behind—and that despite the continuous decrease in global SHDI inequality over time.

## 7 Discussion and Concluding Remarks

The present paper uses the Subnational Human Development Database ([Smits and Permanyer, 2019](#)) to investigate whether, and to what extent, the living conditions across the main subnational units of 163 countries are evolving in a territorially cohesive way. For this purpose, we propose new over- and under-development indicators to complement conventional analyses of inequality, which are often used to study convergence among regions or countries. Specifically, these measures are akin to those commonly used in poverty measurement and allow to identify subnational regions racing ahead or falling behind certain thresholds. Moreover, our approach also allows to construct national-level measures, which may be particularly useful for policy purposes, as they can reflect the share of the population living in over- or under-developed regions and the corresponding degree of over- or under-development.

In our empirical analysis we explore the regional performance in human development: first, relative to the national and then relative to the global average level of human development. Regarding the national-level analysis, our findings suggest that frequently observed reductions of within-country inequality in terms of the HDI are not necessarily accompanied by reductions in under- and over-performance. That is, using the corresponding national HDI as a reference point, not all countries’ subnational HDIs become increasingly similar. All in all, these findings support a more differentiated and cautious interpretation of how countries’ territorial cohesion in terms of the SHDI has evolved over time than what evidenced inequality reductions would suggest.

What about the global distribution of human development across all world countries’ subnational regions? Previous research already documents declines in global inequality of human development ([Jordá and Sarabia, 2015](#), [Permanyer and Smits, 2020](#)). Our results, moreover, suggest that the composition of such declining inequality has shifted over time. In the 1990s, around 23% and 3% of global inequality could be attributed to over- and under-performing regions, respectively. Almost 30 years later, the contributions of over-performing regions fell to 15%, whereas the one of under-performing regions increased to 8%. Thus, while the group of regions racing ahead in human development has declined

over time, the group of under-developed regions did not; as of 2018, it was scattered over some 30 countries.

Remarkably, the set of under-developed regions cuts across national boundaries (i.e. it is not entire countries but rather certain regions within certain countries that are really lagging behind), and the number of individuals living in those areas, circa 308 million has barely declined since 2010 and neither did these regions close the gap with global average human development even partially. Uncovering the existence of this previously undetected—yet non-negligible and time-persistent—pocket of under-development has been possible thanks to the granularity of the SHDI database and the creation of over- and under-development indicators that complement ‘inequality’ and ‘convergence’ approaches. Indeed, while over-/under-performance on the one hand and inequality measures on the other reflect related phenomena, they are intrinsically different, thus providing complementary insights. High inequality in human development does not mechanically imply high regional over- or under-development within a country.

On the methodological side, we stress that all findings reported in the paper are contingent on the choice of the upper and lower cutoffs defining what regions are ‘over-’ or ‘under-developed’, which inherently involves a degree of arbitrariness. It should be reiterated that the choice of the cutoffs is a normative decision, and thus depends on the specific exercise at hand. Moreover, various types of information may enter such considerations, including the ultimate purpose of the measure, the political priorities, and of course the specific distribution of relative performance. All these pieces of information help agreeing on a threshold for how much a region may perform below average in a well-being indicator before being considered as being left behind. These considerations echo the problems and concerns related to the construction of poverty measures, which ultimately depend on the choice of an equally arbitrary poverty line—an issue that does neither invalidate interest nor usefulness of poverty analysis.

What can one conclude from these empirical findings? On the one hand, it seems that many world regions are converging in very basic dimensions of human well-being (i.e. the ones linked to *essential* needs, like survival, basic education, or minimal income). Previous research in similar indicators lends additional support to this as countries’ life expectancy levels tend to increase (despite occasional setbacks) and become increasingly similar globally ([Permanyer and Scholl, 2019](#)) and the number of individuals’ years of schooling continues to increase and become more equally distributed globally ([Jordá and Alonso, 2017](#)) even though developments in regarding global income inequality seem somewhat more complex ([Anand and Segal, 2015](#)). From this perspective, there are several reasons why the SHDI trends around the world since the 1990s can be considered a success story, overall. On the other hand, some of our findings are less inviting for optimism. First, huge pockets of underdevelopment still persist, concealed under national level averages. Second, the overall convergence patterns could suggest that the traditional HDI metric might be less able to dis-



cern the existing differences in living conditions among or within countries. Indeed, there are good reasons to believe that the rosy picture that emerges when using the HDI to assess countries' and regions' socio-economic development can differ dramatically when one expands the focus to incorporate more 'advanced capabilities', i.e. dimensions of human well-being reflecting aspects of life likely to become more important in the near future (or that are currently relevant in high-income settings), like healthy aging, having high-quality and higher education, access to high-level technologies, and so on. Thus, one should be wary of the fact that generalized improvements in basic dimensions of human well-being might co-exist with the emergence of further layers of inequality in more advanced or complex dimensions.

The results shown in this paper document SHDI trends from the late 20<sup>th</sup> century up to 2018. Over this period the SHDI increased in most areas of the world following a rather smooth and monotonic path. The outbreak of the coronavirus pandemic in 2020 might put an end to these trends in a dramatic way. The unprecedented crisis unleashed by the pandemic is strongly affecting each of the HDI's dimensions: for income estimates suggest the largest contraction in economic activity since the Great Depression (e.g., [World Bank, 2020](#)); for health it is anticipated to reduce life expectancy through several different channels (e.g., [Marois et al., 2020](#), [Trias-Llimós et al., 2020](#)); and for education increasing out-of-school rates around the world are already materializing, which are also expected to affect quality-adjusted years of schooling ([Azevedo et al., 2020](#)). The implications that these changes will have on the distribution of human development across and within countries is likely to be harsh—yet still unknown (e.g., [UNDP, 2020](#)).

## References

- Anand, S. and Segal, P., 2015. The global distribution of income, in: *Handbook of Income Distribution*, Elsevier, 937–979.
- Ayala, L., Martín-Román, J., and Vicente, J., 2020. The contribution of the spatial dimension to inequality: A counterfactual analysis for OECD countries, *Papers in Regional Science*, forthcoming.
- Azevedo, J.P., Hasan, A., Goldemberg, D., Iqbal, S.A., and Geven, K., 2020. Simulating the potential impacts of Covid-19 school closures on schooling and learning outcomes: a set of global estimates, Conference edition, The World Bank Group Education.
- Ballas, D., Dorling, D., and Hennig, B., 2017. Analysing the regional geography of poverty, austerity and inequality in europe: a human cartographic perspective, *Regional Studies*, 51 (1), 174–185.

- Bose, A., Chakravarty, S.R., and D'Ambrosio, C., 2014. Richness orderings, *The Journal of Economic Inequality*, 12 (1), 5–22.
- Buhaug, H., Gleditsch, K.S., Holtermann, H., Østby, G., and Tollefsen, A.F., 2011. It's the local economy, stupid! geographic wealth dispersion and conflict outbreak location, *Journal of Conflict Resolution*, 55 (5), 814–840.
- Ceriani, L. and Verme, P., 2015. Individual diversity and the Gini decomposition, *Social Indicators Research*, 121 (3), 637–646.
- Cowell, F., 2011. *Measuring Inequality*, Oxford University Press.
- Decancq, K., Decoster, A., and Schokkaert, E., 2009. The evolution of world inequality in well-being, *World Development*, 37 (1), 11–25.
- Deiwijs, C., Cederman, L.E., and Gleditsch, K.S., 2012. Inequality and conflict in federations, *Journal of Peace Research*, 49 (2), 289–304.
- Durlauf, S.N., Johnson, P.A., and Temple, J.R.W., 2009. The econometrics of convergence, in: *Palgrave Handbook of Econometrics*, Palgrave Macmillan UK, 1087–1118.
- Easterlin, R.A., 2000. The worldwide standard of living since 1800, *Journal of Economic Perspectives*, 14 (1), 7–26.
- Edwards, R.D., 2011. Changes in world inequality in length of life: 1970–2000, *Population and Development Review*, 37 (3), 499–528.
- Ezcurra, R. and Palacios, D., 2016. Terrorism and spatial disparities: Does interregional inequality matter?, *European Journal of Political Economy*, 42, 60–74.
- Ezcurra, R. and Rodríguez-Pose, A., 2017. Does ethnic segregation matter for spatial inequality?, *Journal of Economic Geography*, 17 (6), 1149–1178.
- Foster, J., Greer, J., and Thorbecke, E., 1984. A class of decomposable poverty measures, *Econometrica*, 52 (3), 761–66.
- Foster, J.E., Lopez-Calva, L.F., and Szekely, M., 2005. Measuring the distribution of human development: methodology and an application to mexico, *Journal of Human Development*, 6 (1), 5–25.
- Grimm, M., Harttgen, K., Klasen, S., and Misselhorn, M., 2008. A human development index by income groups, *World Development*, 36 (12), 2527–2546.
- Grimm, M., Harttgen, K., Klasen, S., Misselhorn, M., Munzi, T., and Smeeding, T., 2010. Inequality in human development: An empirical assessment of 32 countries, *Social Indicators Research*, 97 (2), 191–211.

- Harttgen, K. and Klasen, S., 2011. A human development index by internal migrational status, *Journal of Human Development and Capabilities*, 12 (3), 393–424.
- Harttgen, K. and Klasen, S., 2012. A household-based human development index, *World Development*, 40 (5), 878–899.
- Iammarino, S., Rodriguez-Pose, A., and Storper, M., 2019. Regional inequality in europe: evidence, theory and policy implications, *Journal of Economic Geography*, 19 (2), 273–298.
- Johnson, P. and Papageorgiou, C., 2020. What remains of cross-country convergence?, *Journal of Economic Literature*, 58 (1), 129–175.
- Jordá, V. and Alonso, J.M., 2017. New estimates on educational attainment using a continuous approach (1970–2010), *World Development*, 90, 281–293.
- Jordá, V. and Sarabia, J.M., 2015. International convergence in well-being indicators, *Social Indicators Research*, 120 (1), 1–27.
- Kanbur, R. and Venables, A.J., 2005. Rising spatial disparities and development, Policy Brief 3, UNU-WIDER, United Nations University, Helsinki, Finland.
- Kendall, M. and Stuart, A., 1958. *The advanced theory of statistics*, London: C. Griffin.
- Lakner, C. and Milanovic, B., 2015. Global income distribution: From the fall of the berlin wall to the great recession, *The World Bank Economic Review*, 30 (2), 203–232.
- Lessmann, C., 2014. Spatial inequality and development — is there an inverted-u relationship?, *Journal of Development Economics*, 106, 35–51.
- Marois, G., Muttarak, R., and Scherbov, S., 2020. Assessing the potential impact of COVID-19 on life expectancy, *PLOS ONE*, 15 (9), e0238678.
- Mazumdar, K., 2003. Do standards of living converge? a cross-country study, *Social Indicators Research*, 64 (1), 29–50.
- Milanovic, B., 2012. Global inequality recalculated and updated: the effect of new PPP estimates on global inequality and 2005 estimates, *The Journal of Economic Inequality*, 10 (1), 1–18.
- Morrisson, C. and Murtin, F., 2009. The century of education, *Journal of Human Capital*, 3 (1), 1–42.
- Morrisson, C. and Murtin, F., 2013. The Kuznets curve of human capital inequality: 1870–2010, *The Journal of Economic Inequality*, 11 (3), 283–301.

- Neumayer, E., 2004. HIV/AIDS and cross-national convergence in life expectancy, *Population and Development Review*, 30 (4), 727–742.
- Peiró-Palomino, J., 2018. Regional well-being in the OECD, *The Journal of Economic Inequality*, 17 (2), 195–218.
- Permanyer, I., 2013. Using census data to explore the spatial distribution of human development, *World Development*, 46, 1–13.
- Permanyer, I., Esteve-Palos, A., Garcia, J., and Mccaa, R., 2015. Human development index-like small area estimates for africa computed from IPUMS-international integrated census microdata, *Journal of Human Development and Capabilities*, 16 (2), 245–271.
- Permanyer, I. and Scholl, N., 2019. Global trends in lifespan inequality: 1950-2015, *PLOS ONE*, 14 (5), e0215742.
- Permanyer, I. and Smits, J., 2020. Inequality in human development across the globe, *Population and Development Review*, 46 (3), 583–601.
- Pinar, M., 2018. Multidimensional well-being and inequality across the european regions with alternative interactions between the well-being dimensions, *Social Indicators Research*, 144 (1), 31–72.
- Rawls, J., 1999. *A Theory of Justice*, Cambridge, Massachusetts: Harvard University Press, revised ed.
- Riley, J., 2001. *Rising Life Expectancy*, Cambridge University Press.
- Sala-i-Martin, X.X., 1996. The classical approach to convergence analysis, *The Economic Journal*, 106 (437), 1019.
- Salas, R., 2002. Multilevel interterritorial convergence and additive multidimensional inequality decomposition, *Social Choice and Welfare*, 19, 207–218.
- Sen, A.K., 1999. *Development as Freedom*, Oxford: Oxford University Press.
- Sen, A.K., 2009. *The Idea of Justice*, London: Penguin.
- Seth, S., 2009. Inequality, interactions, and human development, *Journal of Human Development and Capabilities*, 10 (3), 375–396.
- Smits, J. and Permanyer, I., 2019. The subnational human development database, *Scientific Data*, 6, 190038.
- Smits, J. and Steendijk, R., 2015. The international wealth index (IWI), *Social Indicators Research*, 122 (1), 65–85.

- Trias-Llimós, S., Riffe, T., and Bilal, U., 2020. Monitoring life expectancy levels during the COVID-19 pandemic: Example of the unequal impact of the first wave on Spanish regions, *PLOS ONE*, 15 (11), e0241952.
- UNDP, 2020. Covid-19 and human development: Assessing the crisis, envisioning the recovery, Tech. rep., United Nations Development Programme, New York.
- Veneri, P. and Murin, F., 2018. Where are the highest living standards? Measuring well-being and inclusiveness in OECD regions, *Regional Studies*, 53 (5), 657–666.
- World Bank, 2020. *Global Economic Prospects, June 2020*, Washington, DC: Washington, DC: World Bank.

## A Additional Results

Table A.1: Countries, number of subnational regions and first survey year

Country	first year	# region	Country	first year	# regions	Country	first year	# regions
AFG	1990	8	GIN	1990	8	NOR	1990	7
AGO	1999	18	GMB	1990	8	NPL	1990	5
ALB	1990	12	GNB	2005	9	NZL	1990	15
ARG	1990	11	GNQ	2000	5	PAK	1990	8
ARM	1990	11	GRC	1990	13	PAN	1990	12
AUS	1990	8	GTM	1990	8	PER	1990	6
AUT	1990	9	GUY	1990	10	PHL	1990	17
AZE	1995	9	HND	1990	18	PNG	1990	22
BDI	1990	5	HRV	1990	21	POL	1990	16
BEL	1990	11	HTI	1990	9	PRT	1990	7
BEN	1990	6	HUN	1990	7	PRY	1990	5
BFA	2000	13	IDN	1990	29	PSE	2004	6
BGD	1990	23	IND	1990	36	ROU	1990	8
BGR	1990	6	IRL	1990	8	RUS	1990	8
BIH	2000	5	IRN	1990	30	RWA	1990	5
BLR	1995	6	IRQ	1990	18	SAU	1990	5
BLZ	1990	4	ITA	1990	21	SDN	1990	15
BOL	1990	9	JAM	1990	6	SEN	1990	10
BRA	1990	27	JOR	1990	12	SLE	1990	14
BRB	1990	4	JPN	1990	10	SLV	1990	4
BTN	2005	20	KAZ	1990	6	SOM	2006	18
BWA	1990	10	KEN	1990	8	SRB	1990	4
CAF	1990	6	KGZ	1990	8	SSD	2010	10
CAN	1990	10	KHM	1990	17	STP	1990	4
CHE	1990	7	KIR	2000	5	SUR	2004	5
CHL	1990	13	KOR	1990	7	SVK	1990	4
CHN	1990	31	KWT	1990	3	SVN	1990	12
CIV	1990	10	LAO	1990	17	SWE	1990	8
CMR	1990	10	LBN	2005	5	SWZ	1990	4
COD	1990	11	LBR	1999	15	SYR	1990	14
COG	1990	12	LBY	1990	3	TCO	2000	8
COL	1990	33	LCA	2000	2	TGO	1990	6
COM	2000	3	LSO	1990	10	THA	1990	5
CPV	2000	5	LTU	1990	10	TJK	1990	5
CRI	1990	7	IVA	1990	6	TKM	2010	6
CUB	1990	15	MAR	1990	7	TLS	2002	13
CZE	1990	8	MDA	1990	4	TTO	1990	5
DEU	1990	16	MDG	2000	22	TUN	1990	7
DJI	1995	2	MDV	1995	6	TUR	1990	12
DNK	1990	5	MEX	1990	32	TZA	1990	25
DOM	1990	9	MKD	2000	8	UGA	1990	9
DZA	1990	7	MLI	1990	8	UKR	1990	5
ECU	1990	3	MMR	1990	14	URY	1990	7
EGY	1990	22	MNE	2006	3	USA	1990	51
ERI	2005	6	MNG	1990	5	UZB	2000	6
ESP	1990	19	MOZ	1990	11	VEN	1990	24
EST	1990	5	MRT	1990	12	VNM	1990	6
ETH	2000	11	MUS	1990	3	VUT	2005	8
FIN	1990	5	MWI	1990	13	XKO	2010	7
FJI	1990	10	MYS	1990	15	YEM	1990	8
FRA	1990	27	NAM	1990	13	ZAF	1990	9
GAB	1990	10	NER	1990	7	ZMB	1990	9
GBR	1990	12	NGA	2003	37	ZWE	1990	10
GEO	2000	10	NIC	1990	3			
GHA	1990	10	NLD	1990	12			

Table A.2: Incidence of over- and underperforming regions  
(a) within country analysis

	under-performance				over-performance		
	(1)	(2)	(3)		(1)	(2)	(3)
	# region-years	# regions	# ctys		# region-years	# regions	# ctys
a=.06	4042	132	63	b=.06	3429	127	69
a=.08	1964	65	35	b=.08	2204	85	51
a=.1	991	33	22	b=.1	1407	58	38
a=.12	490	15	9	b=.12	773	31	18
a=.14	259	12	8	b=.14	503	21	12
Total	49451	1707	163	Total	49451	1707	163

(b) global analysis

	under-performance				over-performance		
	(1)	(2)	(3)		(1)	(2)	(3)
	# region-years	# regions	# ctys		# region-years	# regions	# ctys
a=.15	11186	435	55	b=.15	10717	370	50
a=.2	7547	298	46	b=.2	7250	253	31
a=.25	4129	162	33	b=.25	2950	130	20
a=.3	1819	75	18	b=.3	384	46	8
a=.35	586	24	8	b=.35	0	0	0
Total	49451	1707	163	Total	49451	1707	163

**Notes:** Entire data set contains 163 countries, comprising 1778 regions in total, with 49,451 region-year observations.

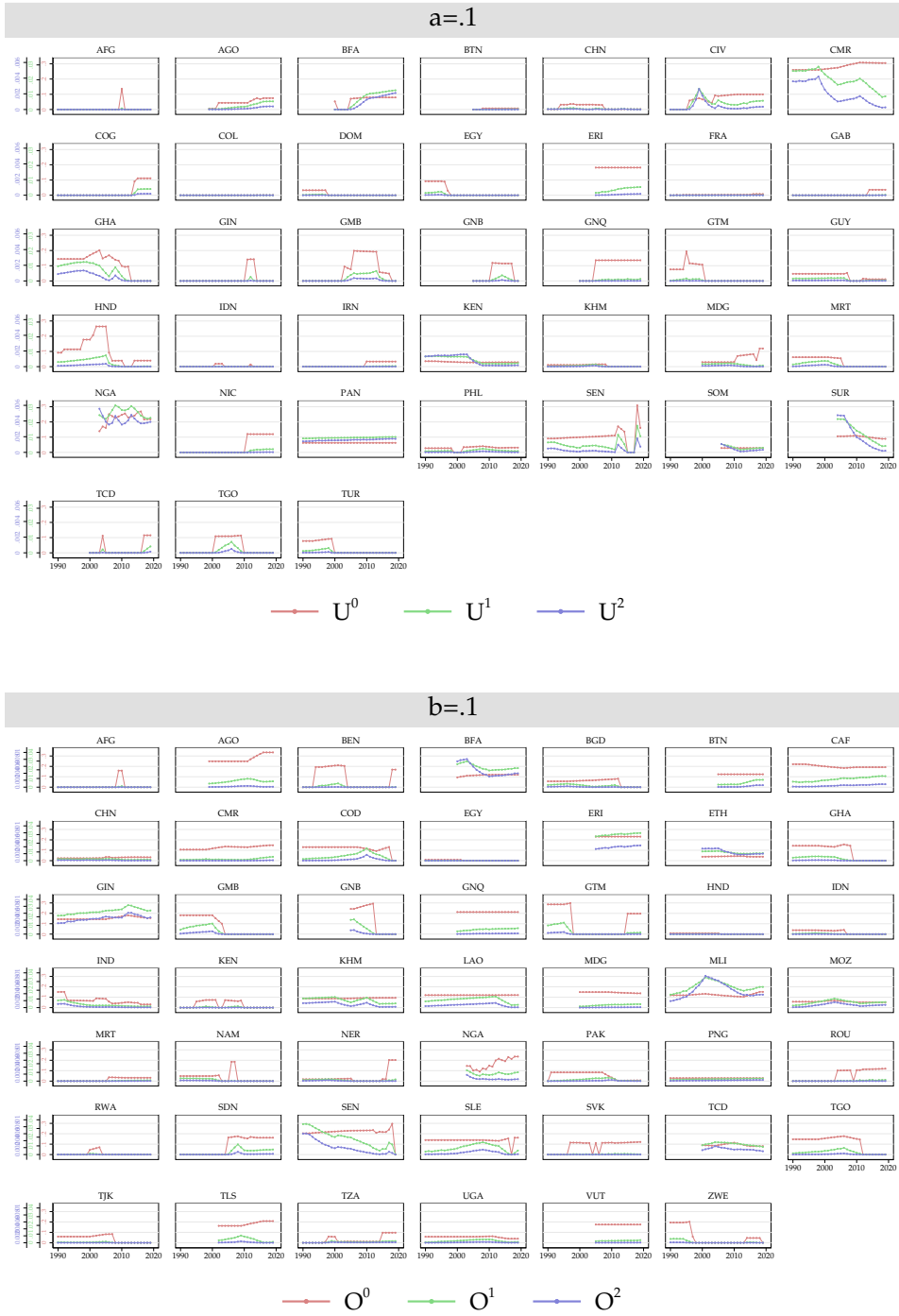
Table A.3: Globally under-performing regions

Country	(1) # underp. regions	(2) share of underp. regions	(3) population share
AFG	1	0.12	0.13
AGO	1	0.06	0.07
BDI	3	0.60	0.62
BEN	1	0.17	0.25
BFA	11	0.85	0.77
CAF	5	0.83	0.81
CIV	2	0.20	0.10
CMR	2	0.20	0.31
COD	5	0.45	0.35
COG	1	0.08	0.07
ERI	5	0.83	0.77
ETH	3	0.27	0.27
GIN	6	0.75	0.70
GMB	5	0.62	0.39
GNB	5	0.56	0.48
KEN	1	0.12	0.03
LBR	11	0.73	0.42
LSO	2	0.20	0.12
MDG	6	0.27	0.24
MLI	7	0.88	0.85
MOZ	6	0.55	0.66
MRT	1	0.08	0.07
MWI	5	0.38	0.38
NER	5	0.71	0.80
NGA	8	0.22	0.28
PNG	1	0.05	0.06
SDN	6	0.40	0.38
SEN	6	0.60	0.43
SLE	11	0.79	0.68
SOM	18	1.00	1.00
SSD	6	0.60	0.63
TCD	7	0.88	0.92
TGO	1	0.17	0.14
YEM	3	0.38	0.42

**Notes:** Data for 2018; underlying threshold  $a = 0.25$ .

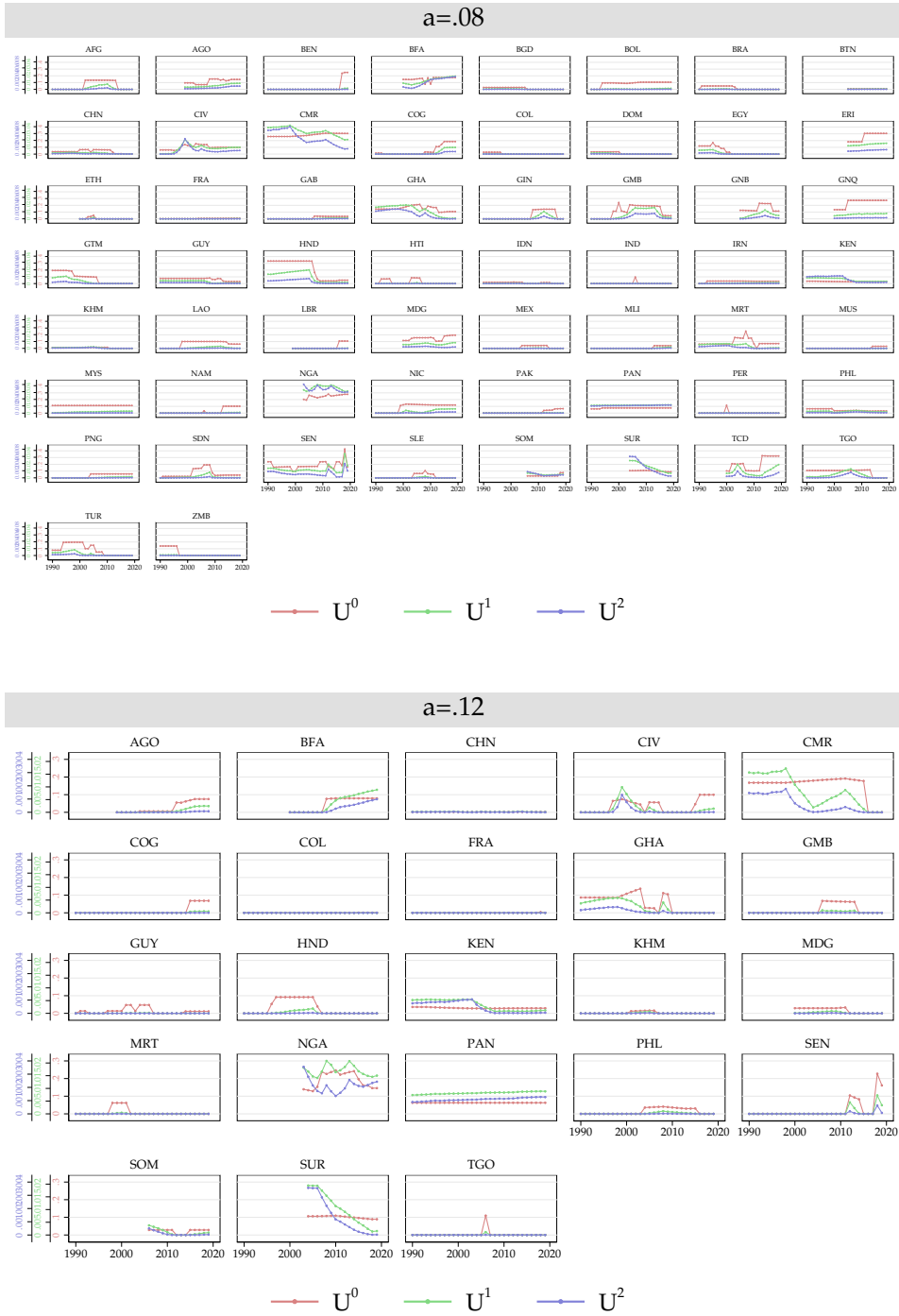


Figure A.1: Over- and Under-performance (preferred threshold)



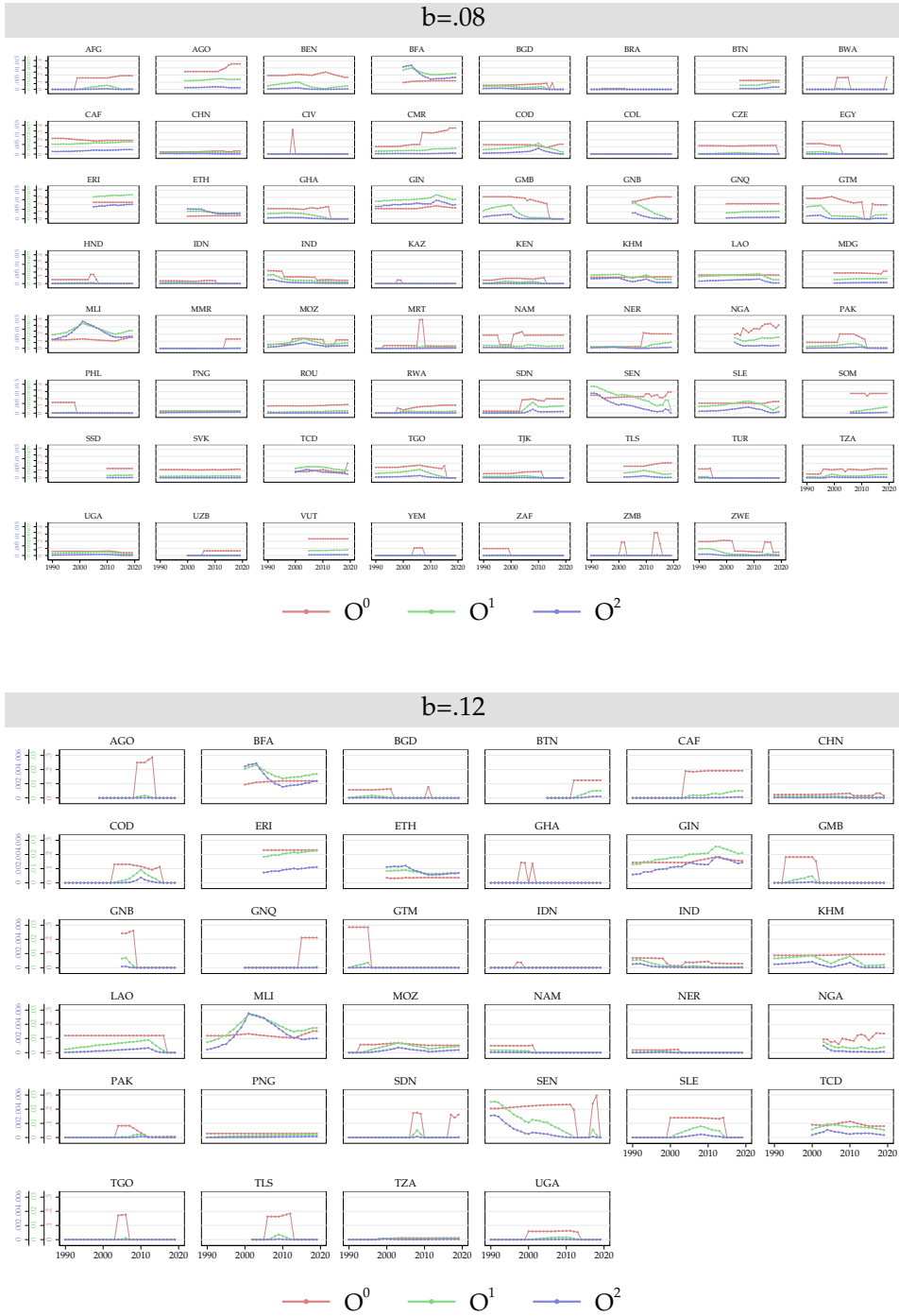
Notes: Figure contains results for all countries having an over- / under-performing region at least once during period of observation.

Figure A.2: Under-performance (alternative thresholds)



Notes: Figure contains results for all countries having an under-performing developed region at least once during period of observation.

Figure A.3: Over-performance (alternative thresholds)



Notes: Figure contains results for all countries having an over-performing region at least once during period of observation.

Figure A.4: Regional contributions to within-inequality—additional results  
(a) Contributions of over- and under-performing regions to within country-inequality

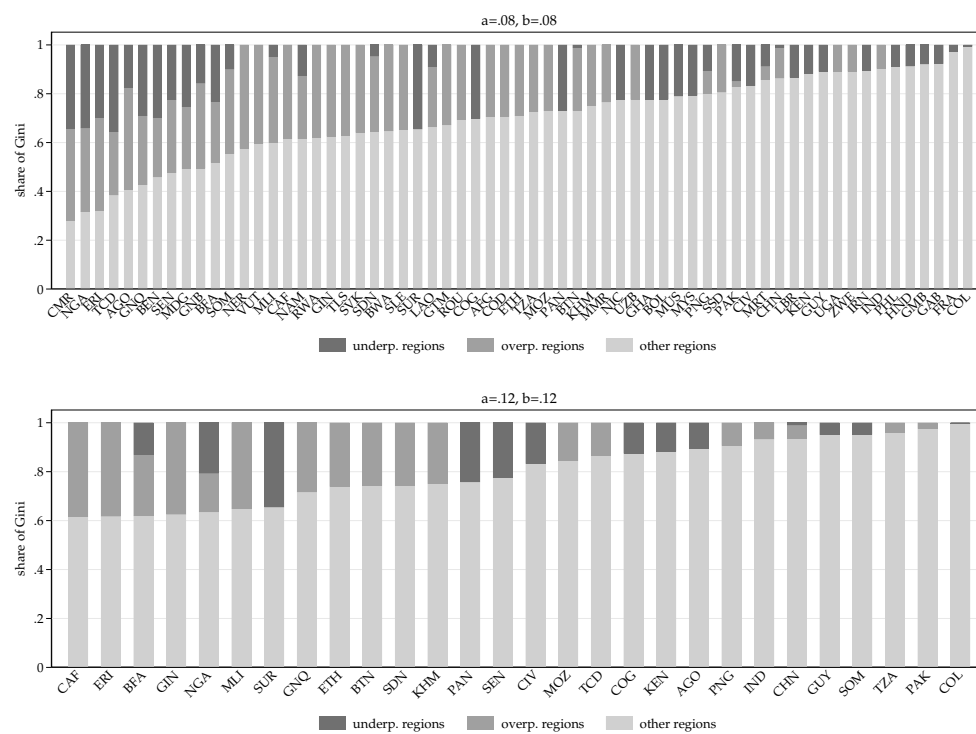


Figure A.5: Regional contributions to inequality over time (selected countries)

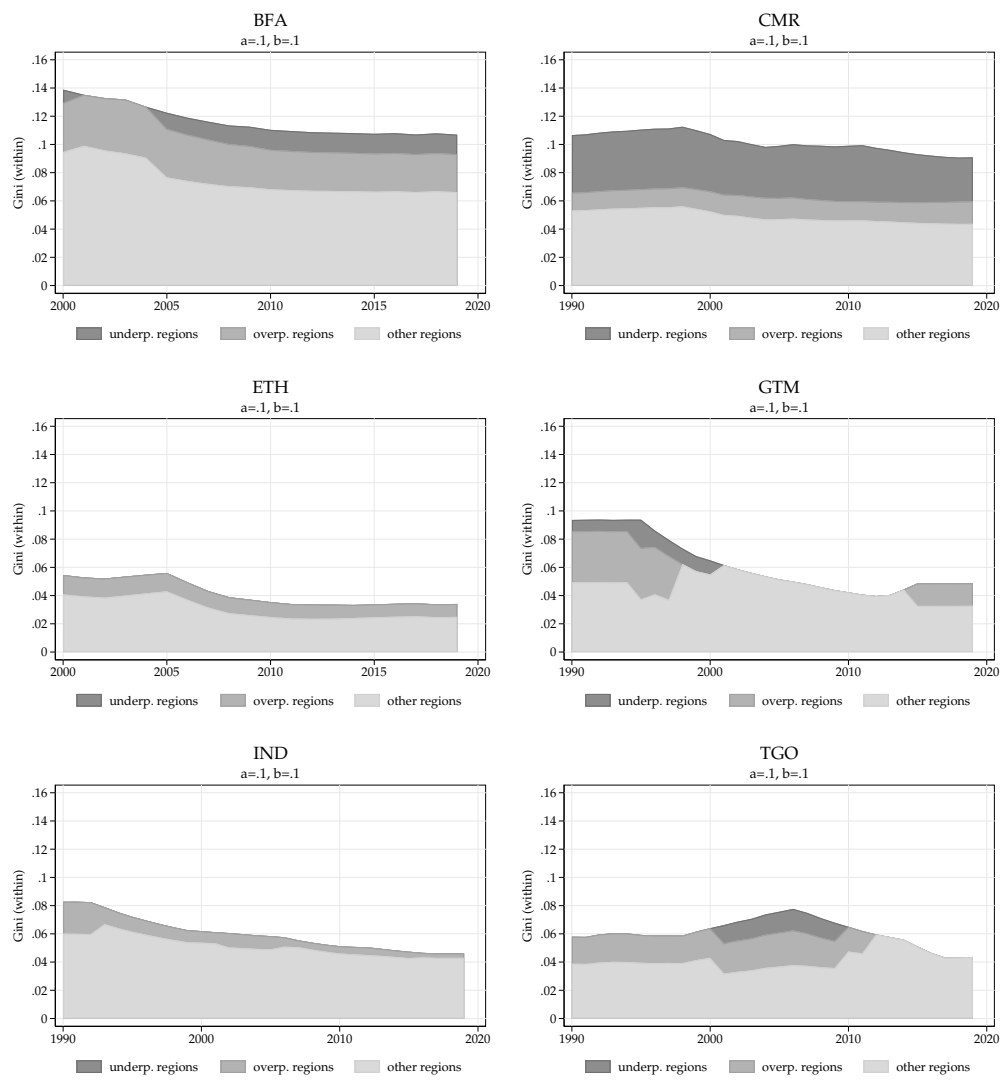
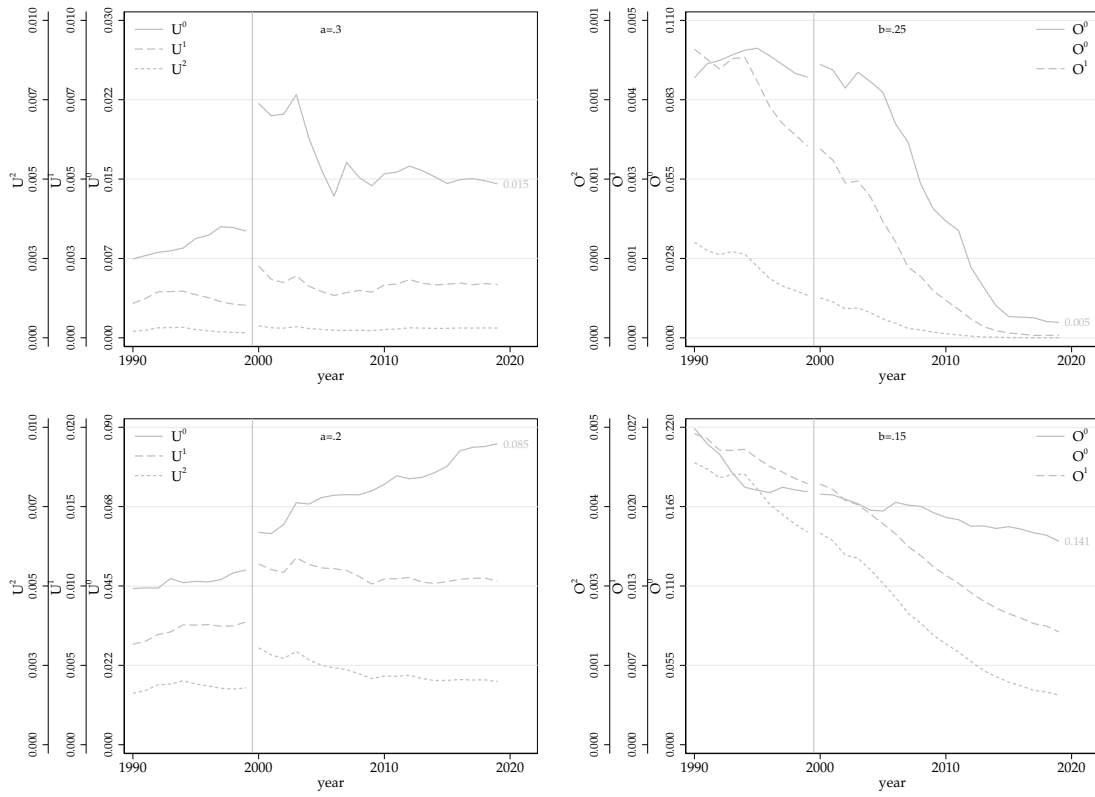


Figure A.6: Global over- and under-performance (alternative thresholds)



Notes: Underlying data is an unbalanced panel. In 2000 data for eleven low-HDI countries becomes available for the first time, thus values before and after 2000 are not directly comparable. Moreover, a few further countries are added in other years, see table A.1 for details.

## B Axioms

Using the notation introduced in this paper, we now list and discuss some of the axioms satisfied by our under-development measure  $U^\gamma$ . Let  $g_i = \max\left\{\frac{z-x_i}{z}, 0\right\}^\gamma$ , and let  $g = (g_1, \dots, g_r) \in [0,1]^r$  be a  $r$ -dimensional vector of under-development gaps. First, we express  $U^\gamma(x,p)$  as a function of the under-development gaps:

$$U^\gamma(x,p) = \sum_i p_i \max\left\{\frac{z-x_i}{z}, 0\right\}^\gamma = \sum_i p_i g_i^\gamma = U^\gamma(g,p)$$

In general, an *under-development measure* is defined as a function  $f:[0,1]^r \times \Delta_r \rightarrow \mathbb{R}_+$ , where  $\Delta_r = \{(p_1, \dots, p_r) \in [0,1]^r \mid \sum_i p_i = 1\}$  is the standard  $r$ -dimensional simplex. For all  $(g,p), (h,q) \in [0,1]^r \times \Delta_r$ ,  $f(g,p) \geq f(h,q)$  is interpreted as indicating that the degree of under-development in  $(g,p)$  is at least as great as the degree of under-development in  $(h,q)$ . We now present an axiomatic characterization of  $f(g,p)$ .

*Continuity* (CON):  $f(g,p)$  is a continuous function.

This axiom ensures that slight measurement errors when measuring the under-development gaps of sub-national units do not result in large errors in our final measure of under-development.

*Normalization* (NOR):  $f(0^{[r]}, p) = 0$ ,  $f(1^{[r]}, p) = 1$ , where  $0^{[r]}, 1^{[r]}$  are  $r$ -dimensional vectors of 0s and 1s, respectively.

This is a standard property ensuring that when there is no sub-national region below the under-development threshold ( $\underline{z}$ ), then the overall under-development measure will take the minimal value of 0. Likewise, whenever all regions attain the lowest development level of 0, then the overall under-development measure will take the maximal value of 1 (this is because, in that case, all under-development gaps  $g_i$  take the maximal value of 1).

*Anonymity* (ANO): Let  $\sigma$  be a one-to-one function from  $\{1, \dots, r\}$  into itself. Then  $f((g_1, \dots, g_r), (p_1, \dots, p_r)) = f((g_{\sigma(1)}, \dots, g_{\sigma(r)}), (p_{\sigma(1)}, \dots, p_{\sigma(r)}))$ .

ANO ensures that the values of our under-development measure do not depend on the labelling of the sub-national regions we are working with.

*Monotonicity* (MON): For all  $g \in [0,1]^r$ , all  $i \in \{1, \dots, r\}$  and all  $g'_i \in [0,1]$ , if  $g'_i > g_i$  then  $f((g_{-i}, g'_i), p) > f(g, p)$ , where  $(g_{-i}, g'_i)$  denotes the vector where the  $i$ -th component of  $g$  ( $g_i$ ) has been replaced by  $g'_i$ .

MON requires that when a subnational region increases its level of under-development, then the corresponding level of overall under-development will also increase.

*Independence (IND)*: For all  $g, h \in [0, 1]^r$ , all  $i \in \{1, \dots, r\}$  and all  $t \geq 0$ , if  $g_i + t \in [0, 1]$ , then  $f((g_{-i}, g_i + t), p) - f((g_{-i}, g_i), p) = f((h_{-i}, g_i + t), p) - f((h_{-i}, g_i), p)$ , where  $(g_{-i}, a)$  denotes the vector where the  $i$ -th component of  $g$  ( $g_i$ ) has been replaced by  $a$ .

IND ensures that our under-development measure is additively separable. Intuitively, this means that the changes in our overall under-development measure ensuing from changes in the development level of a specific region will not be affected by (i.e., are independent of) the development levels of the other regions.

*Uniform Scale Invariance (USI)*: Let  $g, h, g', h' \in [0, 1]^r$ . Suppose  $f(g, p) - f(h, p) = f(g', p) - f(h', p)$ . Then  $f(kg, p) - f(kh, p) = f(kg', p) - f(kh', p)$  for every  $k > 0$ , such that  $kg, kh, kg', kh' \in [0, 1]^r$ , where  $kg = (kg_1, \dots, kg_r)$ .

The intuition behind USI can be explained as follows. Consider two situations with respective under-development gaps  $g$  and  $h$ . Suppose there is an equi-proportionate increase in the under-development levels of all regions in both situations. Then the difference between the levels of under-development in both situations will change by an amount that depends exclusively on the initial difference in under-development and the proportionality factor by which the regional under-development levels change.

*Transfer Axiom (TA)*: Let  $g, h \in [0, 1]^r$ . Suppose that  $g_i = h_i$  for all  $i \in \{1, \dots, r\}$  except for  $i = j$  and  $i = k$ . Assume that  $g_j > g_k$  and  $h_j = g_j + \delta$ ,  $h_k = g_k - \delta$ , for some  $\delta > 0$  in such a way that  $h_j, h_k \in [0, 1]$ . Then  $f(g, p) < f(h, p)$ .

The Transfer Axiom is the cornerstone of inequality measurement. It ensures that after a progressive transfer, our under-development measure should decrease.

We can now introduce the following result axiomatically characterizing our class of under-development measures ( $U^\gamma$ ):

**Theorem**: An underdevelopment measure  $f(g, p)$  satisfies CON, NOR, ANO, MON, IND and USI if and only if, for some strictly positive real number  $\alpha$ ,

$$f(g, p) = \sum_i p_i g_i^\gamma = U^\gamma(g, p)$$

In addition, if  $f(g, p)$  satisfies TA, then  $\gamma > 1$ .



**Proof:** This result is a straightforward adaptation of the paper by Chakraborty, Pattanaik and Xu (2008) to the under-development context we are dealing with here.

**Remark:** The “headcount under-development measure” that one obtains when  $\gamma = 0$  does not satisfy MON, but satisfies CON, NOR, ANO, IND and USI.

The axiomatic characterization of the over-development measure  $O'$  is completely analogous to the one we have just introduced and will not be reproduced here.

## References

Chakraborty, A., Pattanaik, P. and Xu, Y. (2008), “On the mean of the squared deprivation gaps”, *Economic Theory* 34:181-187.