BANCO CENTRAL DE RESERVA DEL PERÚ

Regional Dynamics of Income Inequality in Peru

Luis Eduardo Castillo*

* Banco Central de Reserva del Perú

DT. N°. 2020-004
Serie de Documentos de Trabajo
Working Paper series
Marzo 2020

Regional Dynamics of Income Inequality in Peru*

By Luis Eduardo Castillo†

This Draft: March 2020

This paper discusses the dynamics of income inequality across regions in Peru between 2007 and 2017. Aiming to fill a gap in the usual inequality diagnosis, the article starts by describing the trends in income inequality for each region, and then focuses on (i) identifying the fraction of aggregate inequality that has been explained by inequality between and within regions, and (ii) quantifying the contributions of demographic and socioeconomic factors to these trends. All measurements are done using data from ENAHO, the National Household Survey of Peru.

As regions in Peru are usually understood through political and geographical (longitudinal) categories, I employ both criteria for the purpose of this discussion. The first finding is that all but two political regions (Loreto and Madre de Dios) and all geographical regions in Peru experienced a reduction in inequality between 2007 and 2017 as measured by the Gini coefficient, but the equality gains are highly heterogeneous and seem to have slowed down since 2012. Meanwhile, using Theil indices, I show that most of aggregate inequality in Peru is explained by inequality within regions, although the between component is becoming more relevant just as inequality reduction decelerates. Finally, using counterfactual distributions, I find that the share of adults in the household, labor income, and public monetary transfers have been among the most important drivers of inequality reduction across most regions and in Peru as a whole.

Keywords: Inequality, Gini index, Theil index, Inequality decomposition, Peru

1 Introduction

Although the consensus is that income inequality has grossly declined during the last two decades in Peru and some studies have used quantitative approaches to understand the main factors behind this trend (see, for example, Herrera (2017); Yamada et al. (2016); Azevedo et al. (2013)), little attention have been put towards applying these methods to describe the different dynamics of inequality at the regional level.[1] This clearly represents a problem for fully understanding the inequality phenomenon, given that each region's income distribution has specific characteristics that arise from very different economic and demographic structures, as well as from idiosyncratic income shocks (e.g., redistributive policies). An aggregate decomposition is not then a satisfactory explanation of the divergent trends among regions.

*The views expressed in this paper are those of the author and do not necessarily represent the views of the Central Reserve Bank of Peru. I would like to thank Mario Huarancca, Judith Guabloche and the participants at the BCRP seminar for their insightful discussions.

†Central Reserve Bank of Peru. Email: luiseduardo.castillo@bcrp.gob.pe

[1] Although Seminario et al. (2019), Escobal and Ponce (2012), and Gonzales de Olarte (2010) report different trends in income inequality across regions in Peru, the furthest they go in terms of decomposition analysis is to decompose aggregate inequality into within- and between- contributions.

In this regard, the main purpose of this paper is to analyze the regional disparities in income inequality evolution in Peru between 2007 and 2017. As regions are usually difined through political and geographical categories, I employ both definitions for this analysis.

Specifically, the paper will focus on answering the following two questions:

1. What fraction of aggregate income inequality in Peru during the aforementioned period has been explained by intra- and inter- regional inequality?
2. What are the main sources of the observed changes in income inequality in each region in Peru and how much have they contributed relative to each other?

To accomplish the first task, I employ two of Theil's Generalized Entropy indices for inequality decomposition. These measures, as stated by Fields (2001), are strongly- Lorenz-consistent (just as the widely-used Gini coefficient) but have the advantage that the aggregate value of inequality can be directly decomposed into a weighted average of (i) the inequality within subgroups, and (ii) the inequality between subgroups. The two indices used for this analysis are Theil's L index, in which the individual weights are given by the share of the population, and Theil's T index, in which the weights are the income shares of each region.

Meanwhile, the inequality decomposition at the regional level is made following the strategy of Azevedo et al. (2013). This technique builds up on an accounting structure in which household per capita income is expressed as a function of demographic characteristics and of labor and non-labor income. The strategy consists in creating counterfactual income distributions by replacing *one-by-one* the observed value of each indicator in period 1 with the value of the same indicator in period 0. The inequality measure of the resulting distributions are then interpreted as the levels of inequality that would have prevailed if only those factors had not changed.[2] This process is repeated for every possible decomposition path to allow us to get an average estimate of the contribution of each characteristic to the observed distributional changes, which are known as the Shapley-Shorrocks values.[3] Although the results do not allow for the identification of casual effects, they are useful to identify empirical regularities and recognize the most important elements in inequality evolution from a statistical standpoint.

For the purpose of this research, the socio-demographic characteristics that are being analyzed are the share of adults and share of employed adults in the household (employment ratio). Household income per capita is divided into labor and non-labor income, and non-labor income per capita is further divided into three components: public monetary transfers, rental gains and other non-labor income.

The project is relevant for the Peruvian context because inequality has traditionally been recognized as a pervasive feature of its society, and the aggregate gains in equality may be hiding large heterogeneities that would potentially lead to negative outcomes at the economic, social and political level if ignored. Addressing the divergences in inequality evolution at the regional dimension would effectively identify the regions that are being left behind and could potentially ignite a new interest towards creating a more comprehensive narrative of the income distribution in Peru. Furthermore, the identification of key factors behind inequality trends must be considered when designing policies to curb it.

[2] E.g., the difference between the inequality measure of the observed distribution in period 1 and the one of the counterfactual distribution created by plugging in the period 0 values of labor income is equal to the contribution of this factor to the variation in inequality.

[3] Ibid.

The first finding is that inequality has decreased on aggregate in Peru, but that this trend has slowed down since 2012. At the regional level, although this story is true for almost all regions, there are high variations accross them. When decomposing the aggregate inequality figure into between- and within-regional components, the analysis shows a worrisome result: the between factor has steadily increased since 2012 both when using political and geographical regions as basis for the regional classification. This means that the divergence in income between regions has become more relevant to explain aggregate inequality, just as gains in equality has become smaller. Nonetheless, the between component is still less than 15 percent.

Meanwhile, the quantitative decomposition of inequality shows that, when taking into account the whole window of analysis (2007-2017), there are four almost equally important factors that explain the reduction in inequality across most regions and at the national level: fraction of adults in the household, labor income, current public transfers, and other non-labor income (mostly composed of private transfers). These results are positive in the sense that they suggest that the poorest households in each region have benefited from the demographic boom, economic growth and social policies to catch up with richer households.

However, when using a shorter time frame (2012-2017), the direction of the contribution varies. In the last five years of the analysis, although public transfers and the adult ratio are still grossly equalizing forces, labor income ends up increasing inequality in most political regions. This highlights the importance of public policy in curbing inequality when economic and producitivity growth is slower, but in the longer term policy makers should promote the access to productive jobs among the poorest households to keep doing so.[4]

The remainder of this paper is arranged as follows. Section 2 presents the literature review, giving a brief analysis of the literature on income inequality in Peru and its regions. Section 3 describes the ENAHO, the household survey from which the data come from. Then, Section 4 starts to analyze the inequality phenomenon in Peru by computing inequality measures, and by showing the within and between-region decomposition of inequality (i.e., the answer to the first research question). Next, Section 5 presents the contributions of socio-demographic and economic factors to regional inequality, computed using counterfactual distributions. Finally, Section 6 gives the final remarks.

2 Literature review

2.1 Evolution of inequality in Peru

Since this paper is focused on studying the evolution of inequality, it is worth looking at studies that have measured inequality trends in Peru. In this regard, an appropriate starting point are the official inequality figures published by Instituto Nacional de Estadística e Informática (INEI) (2018).[5] INEI measures inequality both for real household income and expenditure per capita using ENAHO (the same household survey used for this paper). In the income dimension, they show that between 2007 and 2017 the Gini coefficient decreased by 7 percentage points, but that the trend markedly changed after 2012. In fact, between 2007 and 2012, the Gini coefficient declined from 50 to 45 percent, while it only got to 43 percent in 2017. A similar trend is corroborated by Herrera (2017).

[4]The deceleration of economic activity and productivity growth in Peru since 2012 is reported in Castillo and Florián (2019).

[5]INEI is the Peruvian National Institute of Statistics and Informatics.

The main critique of using household survey data for measuring inequality is that the upper end of the income distribution may be underrepresented (Yamada and Castro, 2006). This limitation is in fact later discussed in Section 3. There have been attempts in the literature to find more precise figures both by (i) computing the Gini index with completely alternative data and (ii) using additional information to correct for the skewness in the household income distribution of the surveys. On the first type, we have Alarco et al. (2019) who find different inequality trends depending on the type of data. For instance, using productive wealth series taken from Credit Suisse, they report an increase in inequality since 2013 (they find a similar trend using bank deposit data, but here it is clear that there is a huge underrepresentation of poorer households due to scarce financial deepness in Peru). Meanwhile, when using data from INEI on wages and mixed incomes, they find a reduction of inequality between 2007 and 2016 of 1,0 and 1,4 percentage points, respectively.

On the second type of studies, Yamada and Castro (2006) assume lognormality of the income and consumption distributions and replace the mean value found in the household survey's data with the mean of the same variables in the official national accounts to compute an "adjusted" Gini coefficient. In 2004, for instance, they find a difference of 21 percentage points between the estimate using household survey data and the adjusted one. Later, Yamada et al. (2016) use a similar technique to study the 2007 – 2014 period but employ a new definition of disposable income that includes subsidies and taxes. They also account for a considerable underestimation of the Gini coefficient with household survey data (between 7 and 16 percentage points depending on the year), but the main takeaway is that the decreasing trend remains. In fact, Yamada et al report a more rapid decrease in inequality than INEI.

Other studies include Mendoza et al. (2011), who use a similar strategy as Yamada and Castro to examine the 1985-2010 period but employ GDP and GNP data for the mean substitution. Meanwhile, Cruz Saco et al. (2018) use an approach that assumes that the ENAHO is representative for the first nine deciles of the income distribution, and that the difference between mean income in the household survey and in the national accounts must be inputted only to the top decile. Their results, together with the others that have been mentioned, are presented in Table 1 below.

Table 1. National Gini estimates from other studies

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
INEI (2018)	0,50	0,48	0,47	0,46	0,45	0,45	0,44	0,44	0,44	0,44	0,43
Alarco et al (2018)											
Productive wealth	0,75	0,73	0,77	0,71	0,82	0,80	0,81	-	-	-	-
Wage income	0,19	0,19	0,21	0,20	0,20	0,20	0,19	0,18	0,18	0,18	-
Mixed income	0,18	0,17	0,19	0,17	0,16	0,15	0,15	0,15	0,17	0,18	-
Cruz-Saco et al (2017)	0,68	0,67	0,65	0,65	0,67	0,66	0,66	0,66	0,67	-	-
Yamada et al (2016)	0,65	0,61	0,58	0,57	0,55	0,53	0,52	0,51	-	-	-
Mendoza et al (2011)											
Correction with GDP	0,64	0,64	0,63	0,60	-	-	-	-	-	-	-
Correction with GNP	0,62	0,62	0,61	0,59	-	-	-	-	-	-	-

Source. Own elaboration.

Now, having examined Gini estimations of inequality in the literature, the next question that is worth asking is if there have been any attempts to explain the downward trend in inequality with quantitative approaches. On this note, Jaramillo and Saavedra (2010), using a counterfactual simulation strategy, show that non-labor income inequality (monetary and in-kind government transfers, and private transfers) was the main factor behind inequality reduction between 1997 and 2006, being even more relevant than labor income (which basically remained the same in the period of analysis). In a later study, Jaramillo and Saavedra (2011) use the Theil's T-index decomposition to analyze the within- and between-group contributions to inequality, where the groups are defined by gender, age, education level of the head of the household, and urban/rural area. They report that aggregate inequality is mostly explained by within-group inequality, and that the highest between-group contributions occurred with education (30 percent) and urban/rural area (20 percent).

Meanwhile, Yamada et al. (2016) decompose their Gini estimations into private and public income sources. They find that private income had a stronger equalizing role in 2007-2011 than in 2011-2014. They report that public transfers were key to reducing inequality during the whole period of analysis but were especially crucial between 2011 and 2014, where they explained around 60 percent of all the change in the Gini coefficient.

Finally, Azevedo et al. (2013) compute the Shapley-Shorrocks contribution of demographic and income sources to inequality for 14 Latin American countries, including Peru (using data from SEDLAC).[6] They reveal that, in Peru, 61 percent of the reduction in the Gini coefficient between 2000 and 2010 was due to labor income, and 27 percent was due to an increase in the share of adults in the households (these two are the top equalizing factors). They further show that the share of employed adults actually increased inequality. Herrera (2017) uses the same methodology as Azevedo et al to explain inequality evolution between 2004 and 2015, but separating public transfers from other types of non-labor income. The author concludes that labor income and public transfers were important factors behind inequality reduction.

2.2 Regional disparities in inequality

All the previous studies aimed at describing and explaining inequality on the aggregate level. But, what about the regional evolution of inequality? As it was mentioned in the introduction, studies of this type are scarcer, and usually focus exclusively on describing the trends instead of understanding the contribution of the factors behind them.

Starting again with the official numbers, INEI (2018) publishes the Gini estimates for real household income per capita at the regional level, using for its classification a geographical criteria.[7] In Table 2, it is shown that, according to their estimates, all regions experienced a decrease in inequality between 2007 and 2017. As with the national trend, the decline was steeper between 2007 and 2012 than in the last five years of analysis. When ranking regions, the coast appears to be the most equal, while there is no clear dominance between the highlands and the jungle. INEI does not publish the indices for political regions.

Regarding the explanation of regional divergence in inequality, Seminario et al. (2019) compute the Gini, Theil and Williamson indices for regional GDP between 1795 and 2017 (the authors reconstruct historic GDP data for political regions in Peru). They show that, in all these measures, regional inequality increased between 2000 and 2017 (even after removing Lima from the computations). They also employ the Theil index to decompose the aggregate inequality measure into within- and between-regions contributions, dividing Peru into three regions for this exercise: northern, southern and central. They find that the "within" component explained around 51 percent of aggregate inequality in 2016.

On a similar note, Escobal and Ponce (2012) use income data to compute within- and

[6] Their counterfactual distribution strategy is the same as the one in this paper (see Section 5).

[7] Their classification, however, is not identical to the one used in this paper. In particular, I take the Lima Metropolitan Area out of the coast due to its demographic and economic relevance (see Section 4).

Table 2. Regional Gini estimates from INEI (2018)

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Geographical region											
Coast	0,46	0,42	0,43	0,42	0,41	0,41	0,40	0,40	0,40	0,40	0,40
Highlands	0,52	0,52	0,49	0,48	0,49	0,48	0,47	0,46	0,45	0,46	0,45
Jungle	0,49	0,48	0,49	0,46	0,46	0,46	0,47	0,45	0,46	0,45	0,45
Lima Metropolitan Area	0,46	0,43	0,44	0,43	0,42	0,41	0,41	0,40	0,40	0,41	0,40
Peru	0,50	0,48	0,47	0,46	0,45	0,45	0,44	0,44	0,44	0,44	0,43

Source. Own elaboration from Instituto Nacional de Estadística e Informática (INEI) (2018).

between-regions contributions, using the same classification as INEI (2018). They find that the between contribution was about 12 percent in 2007. Finally, Gonzales de Olarte (2010) reports the Gini coefficient between 2004 and 2007 for the 25 political regions, and shows that the evolution has been largely heterogeneous. None of these studies addresses regional inequality between 2007 and 2017.

3 Data

The data used throughout this paper came from the "Encuesta Nacional de Hogares sobre Condiciones de Vida y Pobreza" (ENAHO), a household survey taken annually in Peru since 2004.[8] The survey's purpose is to shed light on households' living conditions.[9] Although there has been updates and revisions of the survey to include new questions or to improve its design, most of the essential structure has remained intact since 2007. The variables needed for the present analysis are presented in Table 3 below.

Table 3. Description of relevant variables in ENAHO

Variable	Description
MIEPERHO	Total number of members of the household[a]
P208A	Age of the individual
P204	Dichotomic variable indicating if the individual is member of the household
PERCEPHO	Total number of people receiving any type of income (i.e. adults who perceive income)
INGMO1HD	Gross monetary income of the household[b]
INGHOG1D	Gross total income of the household (includes income in the form of goods)[c]
INGTPU02	Income from current public transfers
INGTRAHD	Income from all current domestic transfers
INGRENHD	Rental income
OCU500	Categorical variable for employment status

Note. All monetary variables are expressed in real terms with the deflators in the module "Sumaria".

[a]It excludes domestic workers and individuals subleasing a room in the household.

[b]This variable includes labor income, domestic and foreign current transfer (INGTRAHD and INGTEXHD, respectively), and rental income (INGRENHD). We then need to create a new variable that just comprises labor income by subtracting the other variables.

[c]The strategies to monetarize goods are published in the Technical Note of the ENAHO.

[8]The ENAHO provides both cross-section and panel data, and all the results are publicly available through INEI's website (http://iinei.inei.gob.pe/microdatos/).

[9]For instance, in the 2017 questionnaire, the 371 questions covered household characteristics, household member's characteristics, education, health, employment, income, expenditures, participation in social programs, citizen involvement and individual opinions and perceptions on government and living conditions.

All these variables are self-reported (even the number of members of the household) and this may cause some measurement errors. This caveat is particularly relevant for the income variables, which are constructed from multiple other questions regarding specific sources of income. However, for the purpose of this analysis, I will assume that the values are good proxies of the actual number, and so the ordinal and cardinal differences hold.

Another particular issue with using the ENAHO for measuring inequality is that the households in the top of the income distribution may be underrepresented. Such as Alarco et al. (2019) discuss, the richest household in the survey reports an annual income which is probably a small fraction of the actual income of the richest household in the country. Some strategies to overcome this feature have been discussed in the literature review, but they require us to assume specific shapes of the income distribution that would induce new measurement errors. Instead, if we assume that the direction of the trend remains (which is not a bold assumption after revising the results reported in Section 2), then the survey can still give us valuable insights on the evolution of inequality and on the direction of the relative contributions of factors to this dynamic, which are the focus of this paper.

The population from which the sample is taken consists of all the privately-owned households and their inhabitants in rural and urban areas.[10] The sample sizes (in terms of households) in each wave of the ENAHO are shown in Table 4 below.

Table 4. Sample size of ENAHO waves

Wave of ENAHO	Number of Surveyed Households
2007	22 204
2008	21 502
2009	21 753
2010	21 496
2011	24 809
2012	25 091
2013	30 453
2014	30 848
2015	32 188
2016	35 785
2017	34 584

Note. This is the number of households for whom the income variables are published in the "Sumaria" module.

The survey has a stratified, three-staged clustering sample design.[11]. The stratification is made at the population level (8 ranges of population), but the survey is representative also for urban/rural areas, geographical domains, and for the 25 political regions.

4 Overview of income inequality in Peru

This section presents a detailed description of the inequality phenomenon in Peru. In the first subsection, the analysis is focused on the evolution of inequality on the aggregate and

[10]The survey thus excludes individuals living in collective households (e.g. hotels, retirement houses).

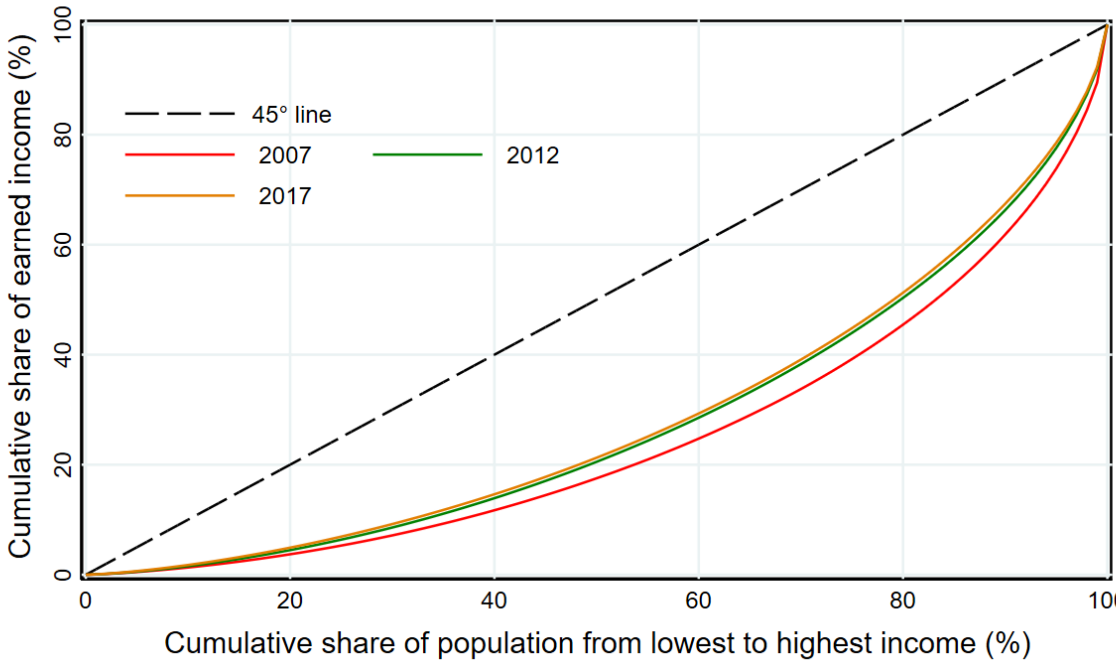
[11]The relevant variable for clustering is CONGLOME (household conglomerate) and for the stratification, ESTRATO (population stratification).

regional level (using the Lorenz curve and Gini coefficient). The second subsection analyses the between-within decomposition of inequality across regions. For all these exercises, and for the remainder of the paper, real gross total household income per capita is employed.[12]

4.1 Evolution of inequality

Figure 1 shows the Lorenz curves for Peru in 2007, 2012 and 2017.[13] The first noticeable feature of the graph is that the curve for the income distribution in 2007 is clearly dominated in the Lorenz sense by both the ones from 2012 and 2017. Domination in the Lorenz sense occurs when a Lorenz curve is never below and somewhere above the other. Meanwhile, when comparing the 2012 and 2017 curves, it is also possible to see that the 2017 curve dominates the one from 2012.[14] Then, the main takeaway of the graph is that any measure of inequality that is Lorenz-consistent (such as the Gini coefficient, entropy measures and Atkinson index) will unanimously yield a decrease in inequality when comparing 2007, 2012 and 2017.

Figure 1. Lorenz curve comparison. Peru, 2007 - 2017.



It is worth mentioning that Lorenz-consistency is a potent property for inequality measures, because it indicates that the inequality ranking with that index will always coincide with the one from Lorenz dominance analysis. This in turn implies that the index encompasses reflexibility, transitivity, anonymity, income homogeneity, population homogeneity, and the transfer principle, which are all desired properties for inequality measures (see Fields (2001)).

The graph above then displays a positive message about income inequality evolution in Peru, at least when considering 5-year variations. However, before proceeding to discuss more about the inequality trends and regional heterogeneities, it is worth asking if these gains in equality are accompanied also by gains in welfare. The Generalized Lorenz curve framework

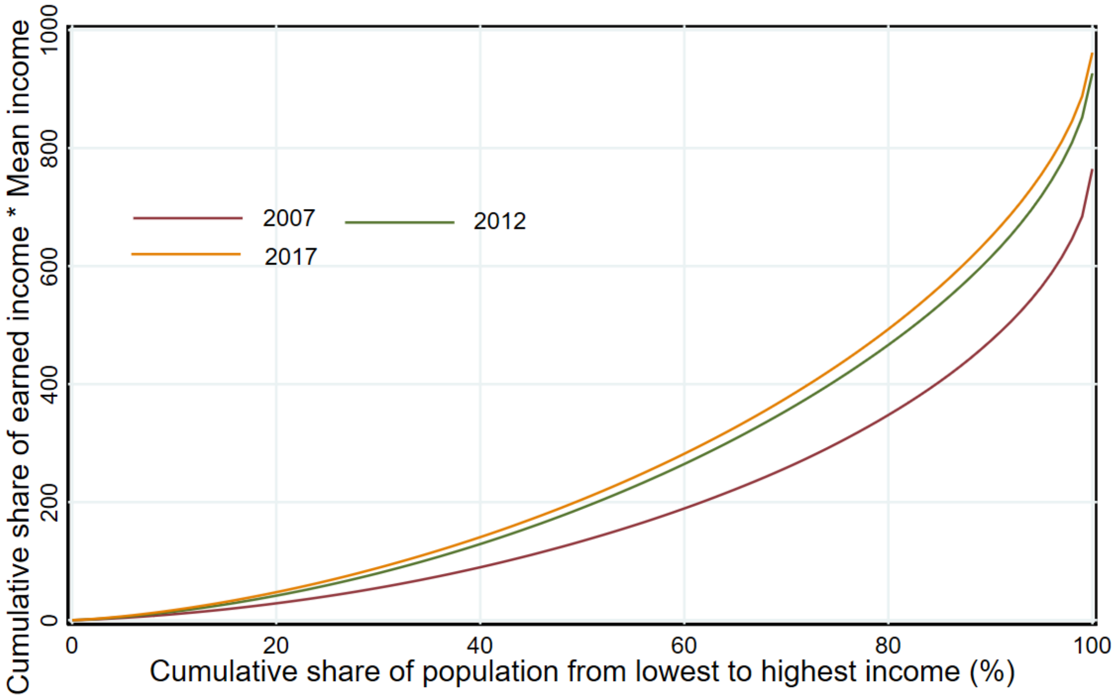
[12]The variable is constructed from INGHOG1D.

[13]All the Lorenz curves and the inequality measures are computed using the Distribuive Analysis Stata Package (DASP). For further information on the package, see Araar and Duclos (2013).

[14]This is harder to distinguish visually because the curves seem to overlap for a significant portion of the range, and the distance between the two curves is shorter than the one between the 2012 and 2007 curves.

is a useful tool to address this enquiry, since it rescales the Y-axis of the Lorenz curve by multiplying it by mean real income. That way, the shape of the curve still gives information about inequality, but the dominance analysis now considers the level of income. As shown in Figure 2, welfare has indeed improved when comparing 2007, 2012 and 2017.

Figure 2. Generalized Lorenz curve comparison. Peru, 2007 - 2017.



Now, returning to the inequality analysis, the initial Lorenz curve graph leaves two particular questions unanswered. On the one hand, it cannot establish cardinal comparisons of inequality between years (although visually it gives some information such as that the gains in equality between 2007 and 2012 should be larger than between 2012 and 2017). On the other hand, since the graph only compares the Lorenz curve from the income distribution in three specific years, there is no guarantee that the inequality decline in between these 5-year gaps has been steady.[15]

Given these considerations, Figure 3 presents the evolution of inequality according to the Gini coefficient, which is a Lorenz-consistent measure widely used in inequality analysis.[16] This index provides both ordinal and cardinal comparisons, and the point estimates allow us to easily detect trends in inequality through time.

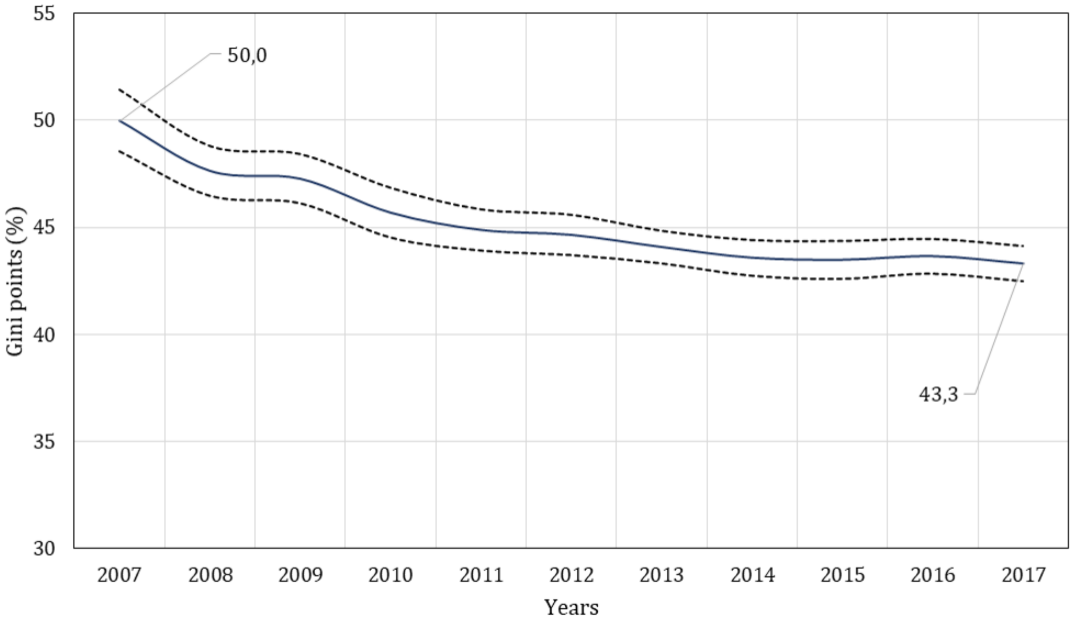
As shown in the graph, inequality in Peru decreased around 7,0 percentage points between 2007 and 2017, while it only diminished one percentage point between 2012 and 2017 (these differences are computed considering exclusively the point estimates of the Gini coefficient). The change in the slope of the curve is also noticeable. Between 2007 and 2012, there is a sharp and steady reduction in the Gini index. Meanwhile, from 2012 onwards, the downward movement has been very smooth, and it appears as if inequality has remained close to being stable when incorporating the confidence intervals in the analysis.

After analyzing the evolution of inequality on aggregate, the next task is to verify that

[15]Although in theory possible, a comparison of Lorenz curves for each of the 11 years being analyzed would be inefficient.

[16]Appendix 7.1 shows the table with the point estimates and confidence intervals.

Figure 3. Gini index evolution. Peru, 2007 - 2017.



the observed trends are the same at the regional level. Before proceeding with the regional analysis, the next set of figures intends to facilitate the discussion by presenting maps of Peru. For the purpose of this discussion, Peru is divided into 25 political regions (Callao is considered a region of its own, and Lima Province, which is politically autonomous, is included into the Lima region to make the classification more comparable to what is usually found in the literature).

Nonetheless, there is an alternative classification of regions based on geographic characteristics. This classification divides the territory into three geographical (longitudinal) regions: coast (the land between the Pacific Ocean and the Andean mountains), highlands (the territory on the Andes), and jungle (the rainforest between the Andean mountains and Brazil).[17] In Panel (b), the yellow, brown and green area correspond to the coast, highlands and jungle, respectively.

Although the political regions are mostly populated in one particular geographical region, almost all of them are located simultaneously in more than one. Given the popularity of this classification in Peruvian academic discussions, I will also run estimations with this regional division of the households. In this analysis, however, I separate the Lima Metropolitan Area (the union of Callao and Lima province) from the coast and considered it a different category due to its relevance both in economic and demographic terms (around a third of the population lives in this area, and most of the economic activity of the country too). This classification will be referred sometimes as "geo-regions" throughout the remainder of this paper.

[17]The Andean mountains cross Peru from the eastern part of Piura to the western part of Amazonas all the way down to the eastern part of Tacna to the eastern part of Puno.

TABLE 5. MAPS OF PERU.

Panel (a) Political map including Callao and Lima province.

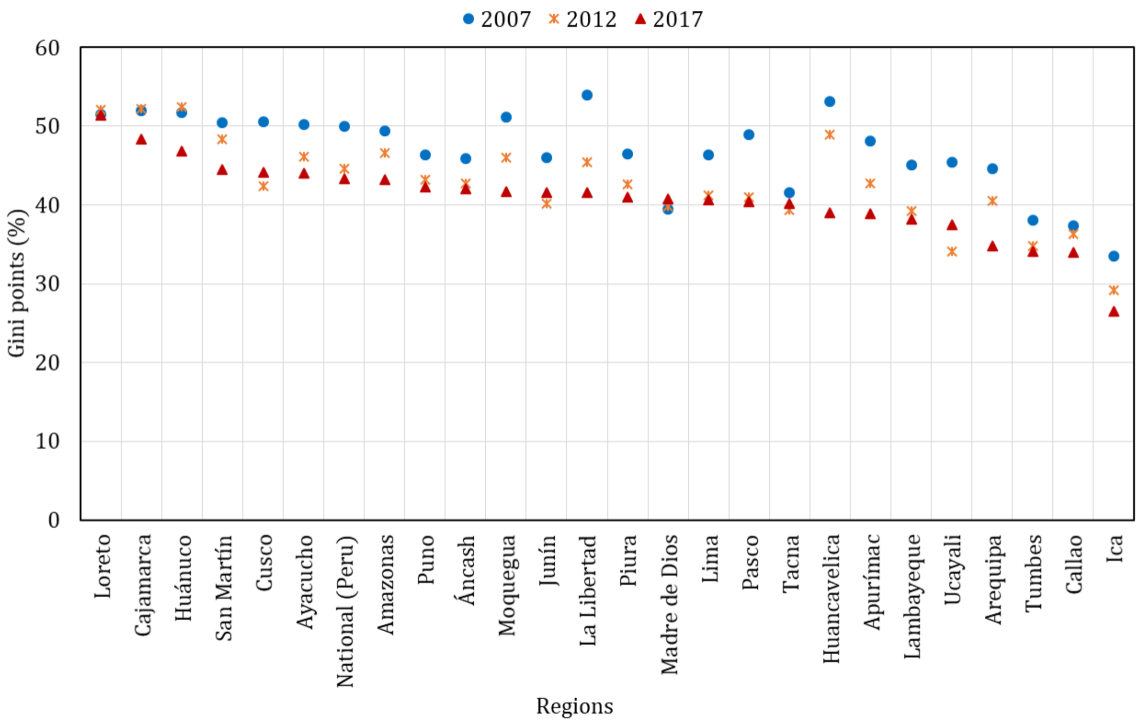
Source. Figure taken from Maps of Worlds.

Panel (b) Political map including geographical division.

Source. Figure taken from Chowell et al. (2009)

Figure 4 now shows the evolution of inequality across political regions for 2007, 2012 and 2017 according to the Gini index. Regions have been ordered from highest to lowest levels of inequality in 2017 to ease the interpretation of the graph. The reduction in inequality between 2007 and 2017 has been experienced by all political regions except Loreto and Madre de Dios. The magnitude of the change, however, varies considerably among the ones who saw gains in equality. There are regions whose reduction has surpassed the 10-percentage points threshold, such as La Libertad and Huancavelica, while other regions have practically experienced no change in the index, like Tacna. This exemplifies the expected variability in inequality evolution across regions.

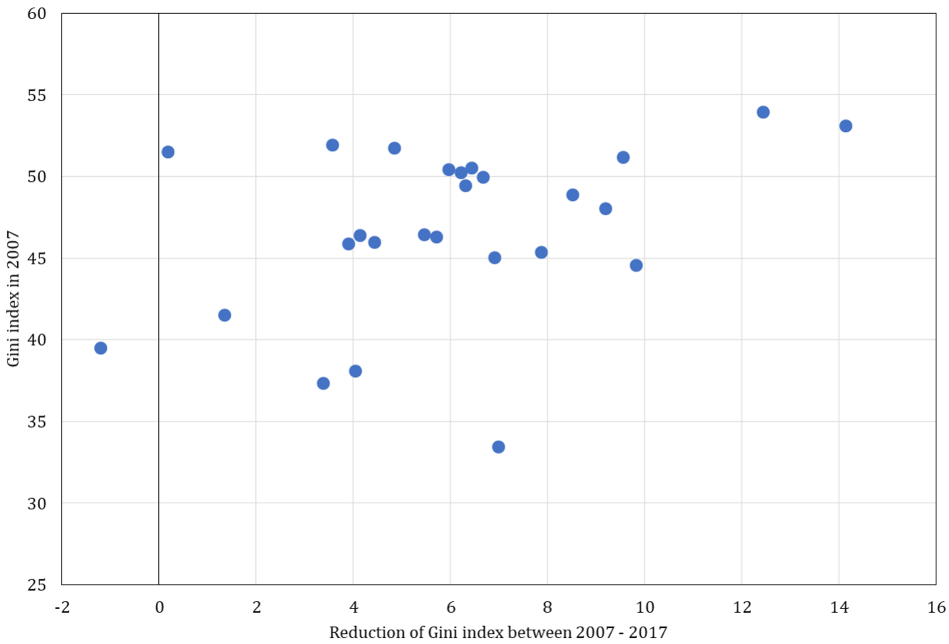
Figure 4. Gini index evolution by political regions. Peru, 2007 - 2017.



Similarly, there is high divergence in the change in the Gini index between 2012 and 2017. For most regions, the reduction of inequality continued, but in a smaller amount than the one observed between 2007 and 2012 (this is visually appreciated by comparing the distance between the blue dot and the yellow asterisk, and the distance between the asterisk and the red triangle). However, there are some regions (Cusco, Junín, Loreto, Tacna and Ucayali), who actually experienced an increase in inequality between the last five years of the analysis. To further appreciate the divergence in trends, Appendix 7.2 shows the point estimates of the Gini coefficient for each region between 2007 and 2017.

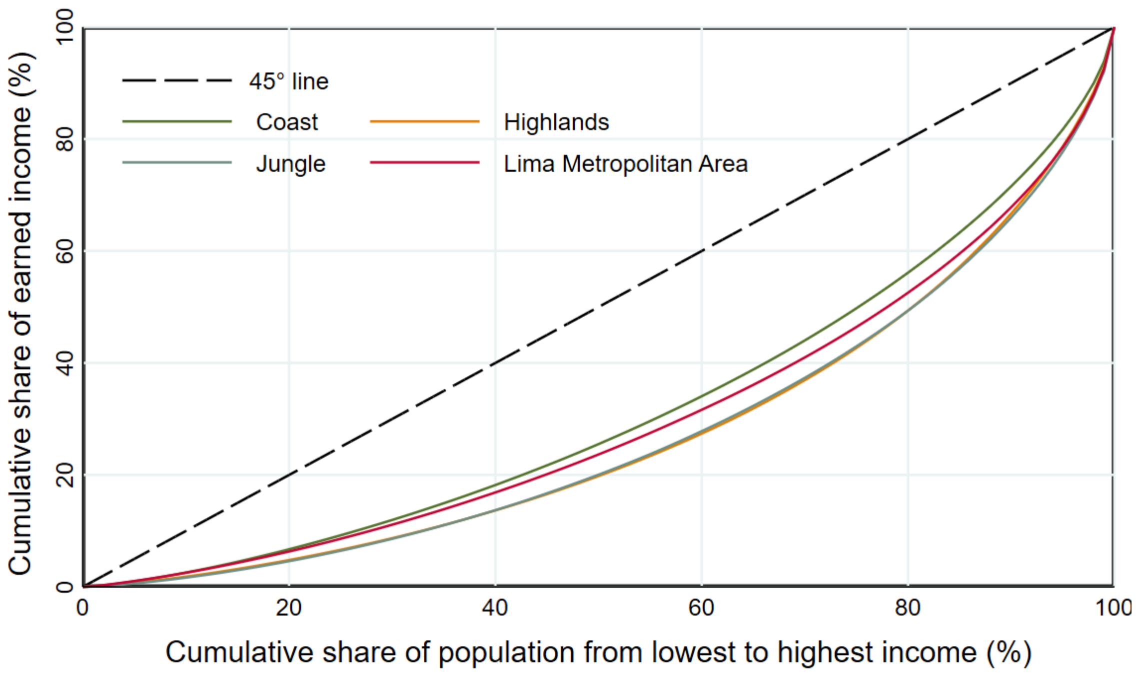
Further studies could relate the observed heterogeneities across regions with other variables. Just as an example, Figure 5 presents a scatter plot of the Gini index in 2007 vs the reduction of the index between 2007 and 2017. There seems to be a positive correlation, hinting at some sort of base effect (and possible convergence).

Figure 5. Variation in Gini index vs index in 2007 by political regions



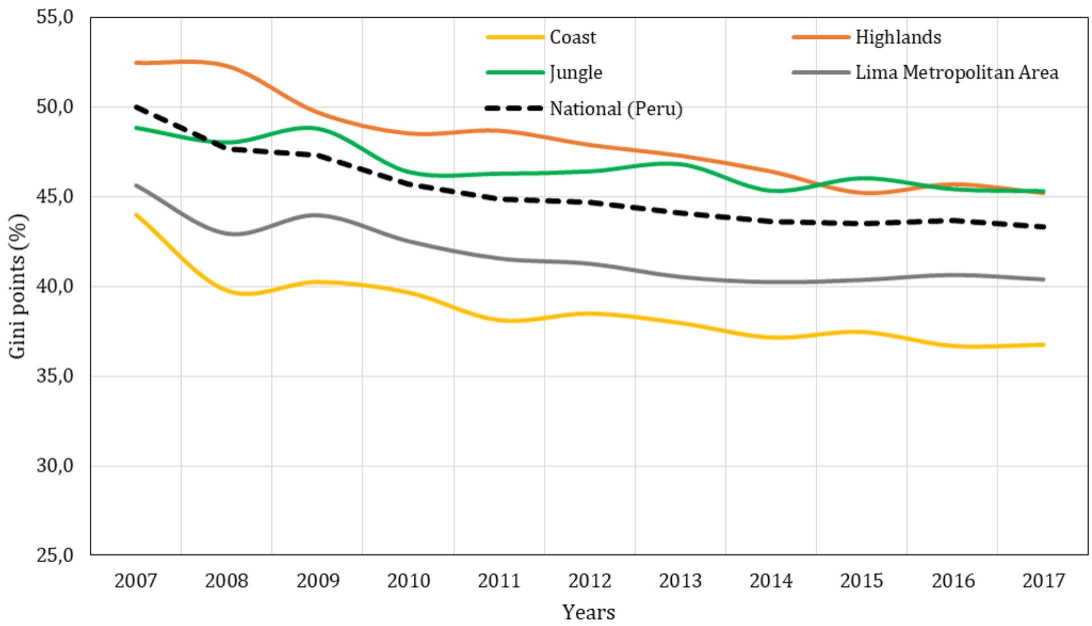
Now, I proceed to make a similar analysis with the geographical regions. Given the smaller number of categories (4 instead of 25 regions), it is now possible to start the analysis with Lorenz-domination between regions. Figure 6 presents the Lorenz curves for geographical regions in 2017. The coast dominates all other regions by the Lorenz criterion, and the Lima Metropolitan Area (LMA) does the same with the two remaining regions. The ranking between highlands and jungle is not possible because the two curves cross each other.

Figure 6. Lorenz curves by geo-regions. Peru, 2017.



To establish if the above dominance relationship has remained the same across time, and if the trends are similar, Figure 7 illustrates the evolution of the Gini index per geographical region (Appendix 7.4 shows these values together with the standard deviation of the estimates).

Figure 7. Gini index evolution by geo-regions. Peru, 2007 - 2017.



It is clear that the previous ranking remains, at least when using the Gini measure. The coast is not only the region with consistently the lowest levels of inequality, but it is also the one who has experienced one of the sharpest decreases between 2007 and 2017 (around 7,3 percentage points). Meanwhile, LMA has lower levels of inequality than the highlands and the jungle, but its inequality reduction has been more gradual than the coast. The picture is particularly interesting because, excluding LMA, the coast is the richest region in per capita terms, followed by the jungle and then the highlands. As the ranking of income level coincides with the ranking of inequality, any measure of well-being should also yield this order.

Furthermore, the previous observed feature of less decline between 2012-2017 than in the 2007-2012 period also holds for this regional classification. Nonetheless, it is particularly interesting the dynamics of the highlands, which continued to experience a sharp decrease in inequality during the last five years of analysis (around 3 percentage points). This got the region closer to the jungle, and in the last couple of years, the difference between the two is statistically not significant. Again, the above graph hints at the importance of considering regional divergence in inequality analysis.

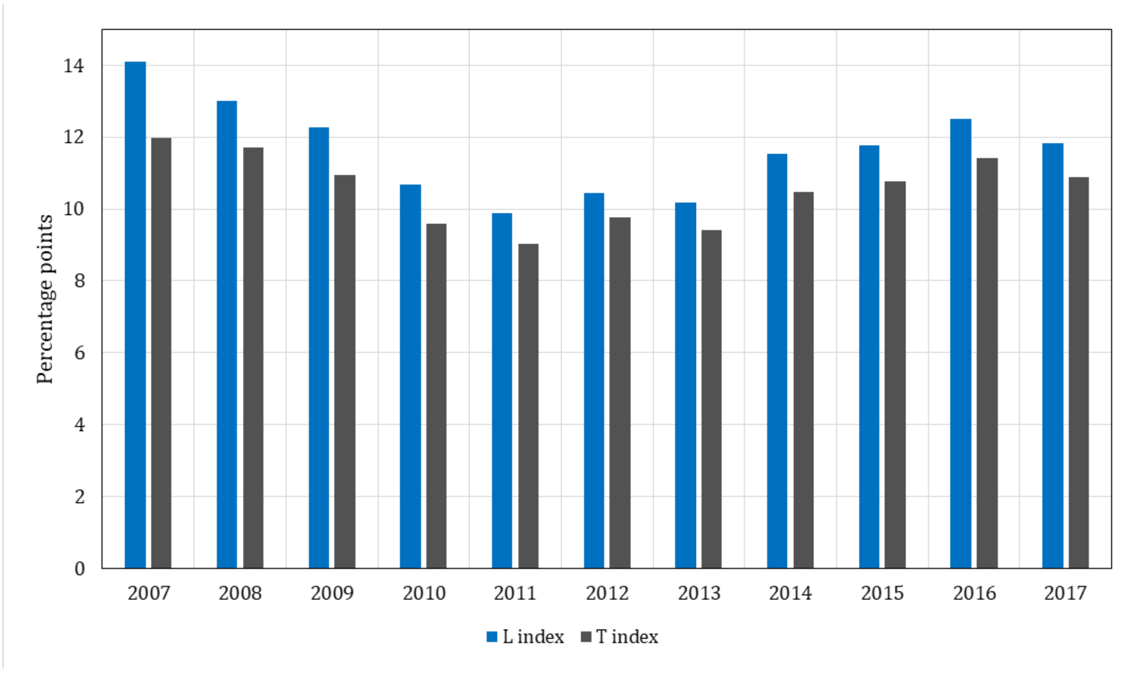
4.2 Between - within regions decomposition of inequality

The between-within decomposition of aggregate inequality is made with Theil's L and T indices, which are Lorenz consistent measures. The indices are variants of the Generalized Entropy measures in which the α parameter (the parameter that defines the sensitivity to the tails of the distribution) takes the value of 0 (L-index) and 1 (T-index). These entropy measures allows us to decompose the inequality measure into a weighted average of (i) the inequality within each region, and (ii) the inequality between regions.[18]

Regarding the specifics of the decomposition procedure, first each region's individual inequality measure is computed. Then, these values are added up with a given weight whose construction depends on the nature of the index. In Theil's L index, the weights are the share of the population. In Theil's T index, they consist of the share of total income. The weighted sum of the individual measures is then called the "within" contribution to inequality. The difference between the aggregate measure of inequality and this "within" factor is defined as the "between" contribution. The contributions are then expressed in relative terms by dividing them by the total value of the aggregate measure.

Figure 8 shows the relative between contribution to inequality after conducting the decomposition across political regions between 2007 and 2017 with the two aforementioned indices. As seen in the graph, the between contribution has remained below 15 percent throughout the period of analysis. This means that most of aggregate inequality is explained by the divergence in income inside each region.

FIGURE 8. BETWEEN-CONTRIBUTIONS. POLITICAL REGIONS, 2007 - 2017.



However, there seems to be a significant increase in the between contribution since 2012. Under both types of measure, 2011 saw the lowest between contribution, but since then it has increased by around 2,0 percentage points. Thus, income divergence between regions has become more important to explain overall inequality in Peru just as the gains in equality in the country started to slow down. This means that not only did most regions stopped moving towards a more equal income distribution since 2012, but that, during this process, some regions were increasingly being left behind in income terms.

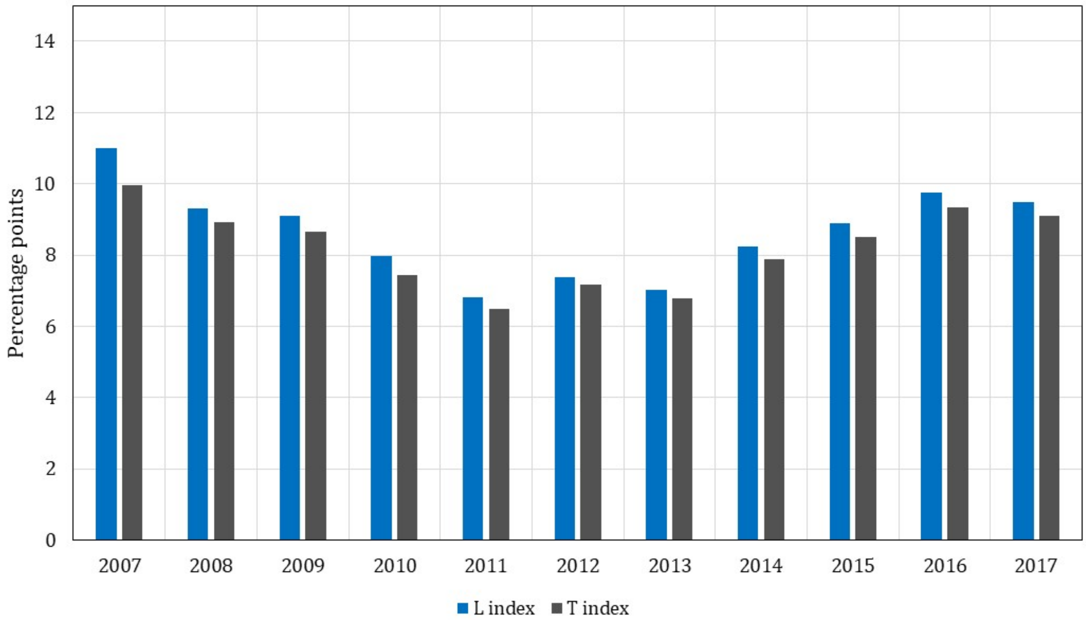
An additional noticeable feature is that the between contribution using the T index is systematically lower than with the L index. As seen in Appendix 7.5 and 7.6, one explanation for this discrepancy is the higher weight of the Lima region to the within component when

[18] This is the decomposition in the sense of Bourguignon (1979). The Gini index does not allow for this type of decomposition, but can only diaggregate the inequality measure with respect to the income sources.

using the income share instead of the population share. As mentioned before, Lima province (the capital), which is included in the Lima region, encompasses significant part of the economic activity of the country. Since Lima region as a whole ranks close to the middle of the regional inequality ranking, it follows that a higher ponderation of its individual index will increase the within contribution of regions (thus lowering the between factor).

Figure 9 presents the same graph as before but with the geographical division.[19] The results are very similar to the ones in the previous analysis. Again, the between contribution is lower with the T index, and this corresponds to the higher weight of LMA (both LMA and Lima have surprisingly similar measures of inequality). The trend in the between contribution is also the same: there has been an increase of this factor since 2011 (around 3,0 percentage points with both indices).

Figure 9. Between contributions. Geo-regions, 2007 - 2017.



5 Factors behind regional income inequality

This section now presents the decomposition of inequality per region into socioeconomic and demographic factors. The first subsection gives a detailed description of the methodology used to create counterfactual distributions and to compute each factor's contribution. The next subsection displays the results and their analysis.

5.1 Methodology

The strategy for inequality decomposition builds on Azevedo et al. (2013), in which counterfactual simulations are used to compute the contribution of each demographic and income component. This method relies on an accounting structure proposed by de Barros et al. (2006), in which per capita household income $\mathsf{Y_{pc}}$ is expressed as follows:

[19] Appendix 7.7 presents the relative contributions of each geographical region per index.

$$\mathsf{Y_{pc}} = \frac{\mathsf{n_A}}{\mathsf{n}}\left[\frac{\mathsf{n_E}}{\mathsf{n_A}}\left(\frac{1}{\mathsf{n_E}}\sum_{\mathsf{i}\in\mathsf{E}}\mathsf{y_i^L}\right) + \frac{1}{\mathsf{n_A}}\sum_{\mathsf{i}\in\mathsf{A}}\mathsf{y_i^{NL}}\right] = \frac{\mathsf{n_A}}{\mathsf{n}}\left[\frac{\mathsf{n_E}}{\mathsf{n_A}}\bar{\mathsf{y}}_\mathsf{E}^\mathsf{L} + \frac{1}{\mathsf{n_A}}\bar{\mathsf{y}}_\mathsf{A}^\mathsf{NL}\right] \tag{1}$$

In the above expression, n represents total members of the household, $\mathsf{n_A}/\mathsf{n}$ is the share of adults in that household, $\mathsf{n_E}/\mathsf{n_A}$ is the share of employed adults, and $\mathsf{y^L}$ and $\mathsf{y^{NL}}$ are labor and non-labor income, respectively. Meanwhile, $\bar{\mathsf{y}}_\mathsf{E}^\mathsf{L}$ is the average labor income of employed adults, and and $\bar{\mathsf{y}}_\mathsf{A}^\mathsf{NL}$ is the average non-labor income of all adults.[20] Thus, this accounting structure recognises that the per capita income of the household is equal to the income earned by adults divided by the number of people living in it, while the income of adults can be divided according to its sources (labor income from adults with a job, and non-labor income).[21]

For the purpose of this research, average non-labor income is divided into smaller components: average current public transfers ($\bar{\mathsf{y}}_\mathsf{A}^\mathsf{PT}$), average rental monetary income ($\bar{\mathsf{y}}_\mathsf{A}^\mathsf{R}$), and the average of other non-labor income ($\bar{\mathsf{y}}_\mathsf{A}^\mathsf{ONL}$).

Since the cumulative density function of households' income $\mathcal{F}$ depends on $\mathsf{Y_{pc}}$, and any measure of inequality θ (e.g. Gini index, Theil's indices) depends on this cumulative density function, then θ could be expressed as follows:

$$\theta = \theta\left(\mathcal{F}\left(\mathsf{Y_{pc}}\right)\right) = \theta\left(\frac{\mathsf{n_A}}{\mathsf{n}}, \frac{\mathsf{n_E}}{\mathsf{n_A}}, \bar{\mathsf{y}}_\mathsf{E}^\mathsf{L}, \bar{\mathsf{y}}_\mathsf{A}^\mathsf{NL}\right) = \theta\left(\frac{\mathsf{n_A}}{\mathsf{n}}, \frac{\mathsf{n_E}}{\mathsf{n_A}}, \bar{\mathsf{y}}_\mathsf{E}^\mathsf{L}, \bar{\mathsf{y}}_\mathsf{A}^\mathsf{PT}, \bar{\mathsf{y}}_\mathsf{A}^\mathsf{R}, \bar{\mathsf{y}}_\mathsf{A}^\mathsf{ONL}\right) \tag{2}$$

Given that the levels of all the indicators above are known in period 0 and period 1, the counterfactual distributions for period 1 are constructed by replacing the observed magnitudes of the indicators in period 0 one at a time. Hence, after plugging in the period 0 level of an indicator, the inequality measure of this new counterfactual distribution can be interpreted as the inequality level that would have prevailed in the absence of a change in that indicator between period 0 and period 1. It is clear then, as stated by Azevedo et al. (2013), that this decomposition strategy *does not identify causal effects*, but instead intends to describe the elements that are quantitatively more important in distributional changes (i.e., it finds empirical regularities in the data).

The measure of inequality θ that will be used for this analysis is the Gini coefficient.

As an example of how this works, given the observed Gini index for period 1, θ_1, and a Gini index constructed from an income distribution where all the variables except average labor income correspond to period 1, $\hat{\theta}_1 = \theta\left(\frac{\mathsf{n_A}}{\mathsf{n}}, \frac{\mathsf{n_E}}{\mathsf{n_A}}, (\bar{\mathsf{y}}_\mathsf{E}^\mathsf{L})_0, \bar{\mathsf{y}}_\mathsf{A}^\mathsf{NL}\right)$, the difference $\theta_1 - \hat{\theta}_1$, would be the contribution of labor income to the change in inequality between period 0 and period 1 for this particular decomposition path (i.e., replacing labor income first). Then, if we replaced the value of the share of adults in period 0 and get a new Gini measure $\hat{\theta}_2$, the difference $\hat{\theta}_1 - \hat{\theta}_2$ would be the contribution of the share of adults to inequality for this particular decomposition path (i.e, replacing labor income first and the share of adults, second).

[20] Notice that the subscript indicates the population from which the average is taken from: A for adults and E for employed adults.

[21] In the Peruvian context, adults would be understood as the individuals 14 years old or above who in theory are the only ones that are able to work (this assumption, however, is not perfect because Peru is known to have high rates of child labor, specially in rural areas, even for South American standards).

The first obvious problem that arises with this method is that, in the absence of panel data, there is no clear way in which to input the values from period 0, since it is difficult to identify which household in period 1 should be the equivalent of a household in a previous year. Azevedo et al suggest addressing this problem by: (i) ordering the households by their household income per capita in both periods, (ii) taking the average value of the indicator for each quantile in period 0, and finally (iii) assigning this average value to the households of the same quantile in period 1. In these computations, 200 quantiles are employed.

The second problem is that the results suffer from path-dependence, meaning that the order in which each characteristic is inputted matters for the calculation of its contribution. Given that there are 6 variables, there are in fact 6! paths for decomposition. To solve for this, Azevedo et al calculate the decomposition across all possible paths, and then compute the average results for each component, which are called the Shapley-Shorrocks estimates.[22]

5.2 Results

Figures 10 and 11 present the contributions of each of the six factors to the gains in equality between 2007 and 2017 for the political regions. Gains in equality are just the negative of the variation in the Gini coefficient, and the results are expressed in this way so that a positive contribution means that the factor contributed to the reduction in inequality.[23]

In Figure 10, the contributions are expressed in percentage points, thus being absolute contributions (i.e., the stacked bars sum up to the total equality gains in Gini points). In Figure 11, the contributions are expressed as a fraction of the total equalities gains, thereby being relative contributions. In this figure, the green and yellow upward arrows are used for positive relative contributions above 10 and between 0 and 10 percent, respectively; while the red and gray downward arrows are displayed for negative relative contributions below -10 and between -10 and 0 percent, respectively.

The figures show common patterns between individual regions, as well as between the regional results and the national decomposition. The first noticeable feature is that the adult ratio has contributed positively to equality in almost all the studied political regions and at the aggregate level. This indicates that the demographic transition in Peru has been a remarkable equalizing force. According to estimates of INEI, the population over 14 years (allowed to work by law), has increased sharply since the 2000s.[24] The present analysis is telling us that the benefits of this demographic boom and contraction of the dependency ratio has been strongly experienced by poorer households in most regions.

However, the positive effect on equality of this demographic boom appears to have not been corresponded by the share of employed adults. In fact, in most regions (and in Peru as a whole), this factor has increased inequality, which means that households at the bottom of the distribution have been less capable of getting jobs relative to richer households.

[22] All the procedure can be done with the Stata ado ADECOMP, developed by Azevedo et al. (2012).

[23] The magnitude of the change in Gini points may not be identical to the one from the previous analysis because the ADECOMP tool drops some observations when constructing the counterfactual distributions.

[24] This can be easily seen in: http://webapp.inei.gob.pe:8080/sirtod-series/.

Figure 10. Absolute contributions to gains in equality (Gini points). Political regions, 2007 - 2017.

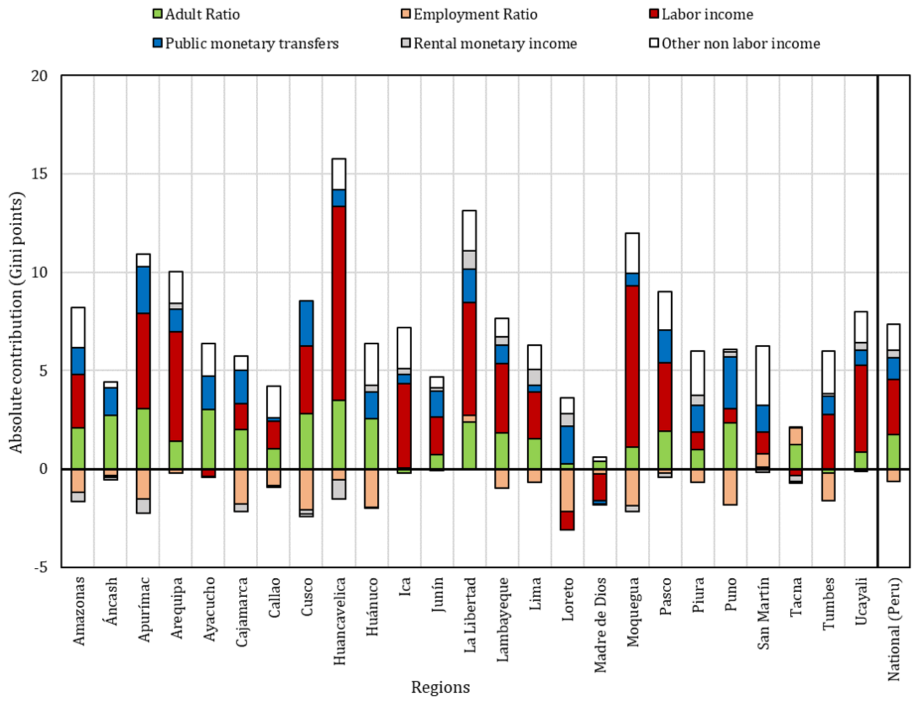


Figure 11. Relative contributions to gains in equality (%). Political regions, 2007 - 2017.

Regions	Adult Ratio	Employment Ratio	Labor income	Public monetary transfers	Rental monetary income	Other non labor income	Gains in equality (Gini points)
Amazonas	32	-18	41	21	-7	31	**7**
Áncash	71	-8	-3	35	-3	8	**4**
Apurímac	35	-18	55	27	-8	7	**9**
Arequipa	14	-2	57	12	3	17	**10**
Ayacucho	51	0	-6	28	-1	28	**6**
Cajamarca	56	-50	37	47	-10	20	**4**
Callao	31	-26	43	5	-2	49	**3**
Cusco	46	-34	56	38	-4	-2	**6**
Huancavelica	24	-4	69	6	-7	11	**14**
Huánuco	59	-45	-1	31	8	49	**4**
Ica	-3	1	61	7	4	30	**7**
Junín	16	-2	41	29	3	13	**5**
La Libertad	18	2	44	13	7	15	**13**
Lambayeque	27	-14	53	14	6	14	**7**
Lima	28	-12	42	6	14	22	**6**
Loreto	51	-417	-174	369	118	153	**1**
Madre de Dios	33	-21	-113	-16	-2	19	**-1**
Moquegua	11	-19	83	6	-3	21	**10**
Pasco	22	-3	41	19	-2	22	**9**
Piura	19	-13	17	26	9	42	**5**
Puno	55	-42	16	61	7	3	**4**
San Martín	1	11	19	22	-3	49	**6**
Tacna	86	60	-23	4	-21	-6	**1**
Tumbes	-5	-32	63	22	3	49	**4**
Ucayali	11	-2	56	10	5	20	**8**
National (Peru)	26	-9	42	16	6	19	**7**

Second, both labor income and other non-labor income (which is mainly constituted of private transfers) have lessened inequality in most regions. In fact, at the national level, labor income contribution is the highest among all factors. This is a jubilant outcome, because it means that economic growth occured in such a way that the labor income and private transfers of poorer households expanded enough in relative terms to reduce inequality.

Finally, from a policy standpoint, the most important result comes from the contribution of monetary public transfers. In all but one political region (Madre de Dios), public monetary transfers were instrumental in the decline of inequality. For instance, in 10 of the 25 regions, public transfers represented more than a quarter of the reduction in the Gini coefficient. This hints at the possibility that social policy had been targeted well enough to have distributional effects. The hint is stronger when seeing that some regions with the strongest relative contribution are among the poorest ones (Apurímac, Ayacucho, Cajamarca and Puno).

Now, regarding the geographical division of the regions, Figure 12 presents the relative contributions to the gains in equality for this classification. In the figure, we see that the main story holds for geographical regions: adult ratio, labor income, other non-monetary income and public monetary transfers have contributed positively to equality, while the adult employment ratio has increased inequality.

Having a smaller set of regions allows us to easily compare the magnitudes of these effects. On the one hand, it is easy to see that labor income has contributed the most in LMA and in the coast, which is not surprising considering that these regions are the richest and most productive ones, and probably the poorer households there have experienced some of the highest income growth in the country. On the other hand, the contribution of public monetary transfers is larger in the highlands and in the jungle, and practically non-existent in LMA. This again hints at the story of good targeting because it means that social policies aimed at curbing inequality are having most effect among the traditionally poorest households.

Figure 12. Absolute contributions to gains in equality (Gini points). Political regions, 2007 - 2017.

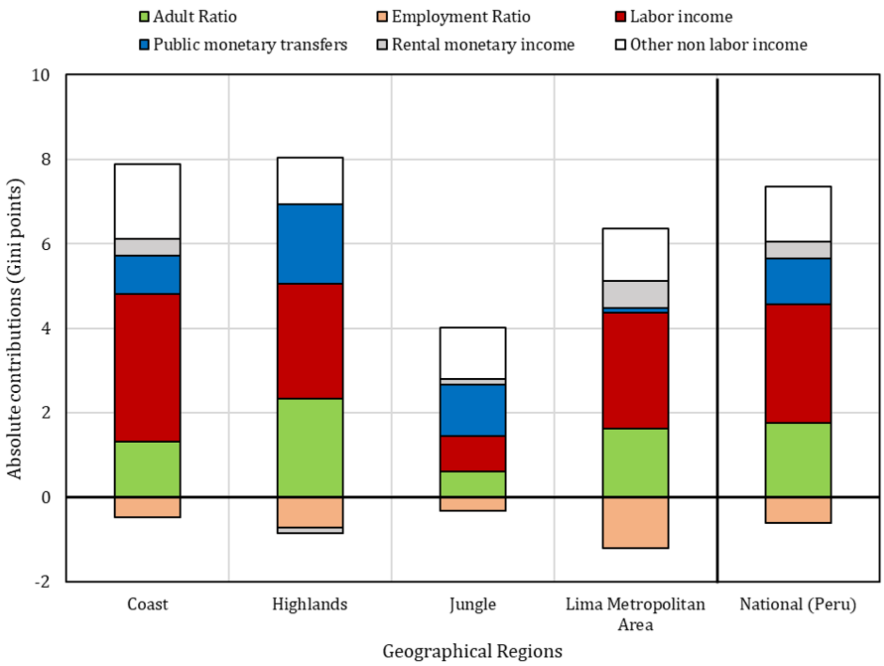


Figure 13. Relative contributions to gains in equality (%). Geographical regions, 2007 - 2017.

	Adult Ratio	Employment Ratio	Labor income	Public monetary transfers	Rental monetary income	Other non labor income	Gains in Equality (Gini points)
Coast	18	-6	47	12	5	24	**11**
Highlands	33	-10	38	26	-2	15	**8**
Jungle	17	-8	23	33	4	33	**5**
LMA	31	-23	53	2	13	24	**6**
National (Peru)	26	-9	42	16	6	19	**7**

The analysis in Section 4 showed that there was a shift in the inequality trend from 2012 onwards. In this regard, it is worth replicating the previous exercise for the 2012-2017 period to see how the contribution of each factor differs when accounting only for the last five years of analysis. Figures 14 and 15 present these results for the political regions.

There is a noteworthy change in the story of the factors when taking into account this time frame. Labor income, which was a prominent equalizing force when comparing 2007 with 2017, has actually increased inequality between 2012 and 2017 in many regions. Moreover, the negative relative magnitude is quite high throughout them. Then, the income growth differential between poorer and richer households must have shortened in this period.

Figure 14. Absolute contributions to gains in equality (Gini points). Political regions, 2012 - 2017.

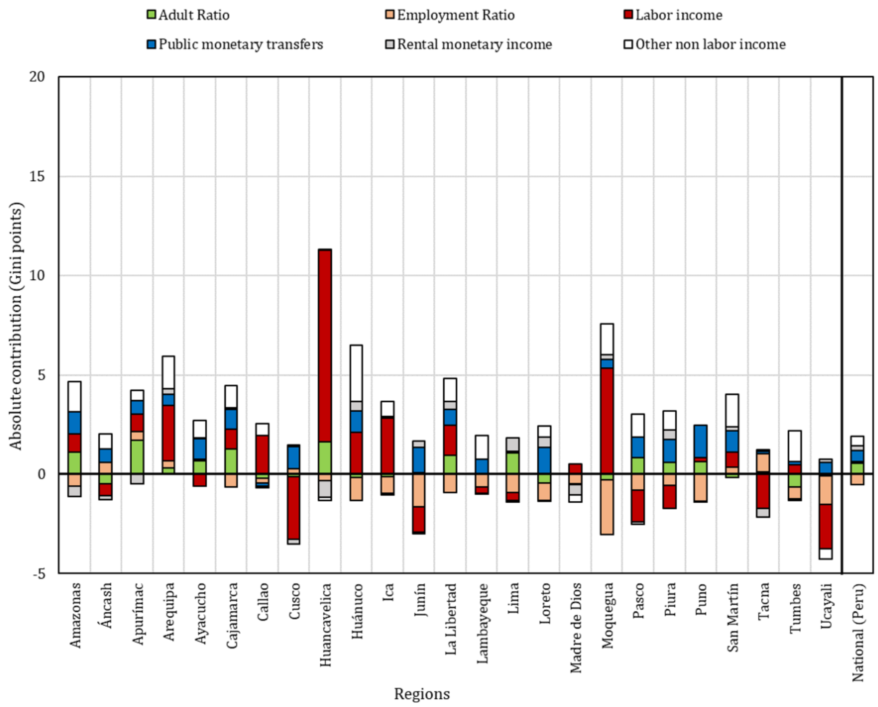
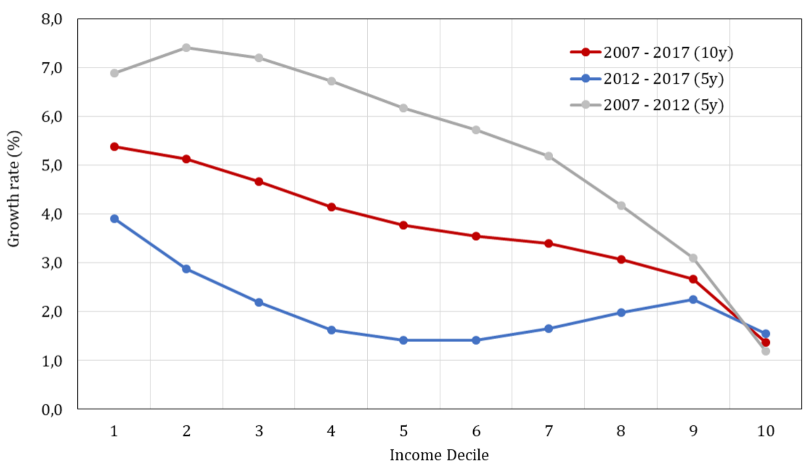


Figure 15. Relative contributions to gains in equality (%). Political regions, 2012 - 2017.

Regions	Adult Ratio	Employment Ratio	Labor income	Public monetary transfers	Rental monetary income	Other non labor income	Gains in Equality (Gini points)
Amazonas	32	17	-25	32	15	43	**4**
Áncash	-65	-79	75	86	-26	-101	**1**
Apurímac	45	-12	-23	17	12	14	**4**
Arequipa	5	-6	-47	-10	-4	28	**6**
Ayacucho	32	-5	29	48	-3	-40	**2**
Cajamarca	34	17	-25	26	-3	29	**4**
Callao	-12	12	-103	8	4	33	**2**
Cusco	-7	13	-153	-54	-12	4	**-2**
Huancavelica	16	3	-96	0	-8	-2	**10**
Huánuco	-3	22	-41	-20	-10	54	**5**
Ica	-4	33	-107	2	-3	29	**3**
Junín	7	-126	-95	97	23	-5	**-1**
La Libertad	25	23	-38	20	-11	-30	**4**
Lambayeque	0	64	31	75	1	-120	**1**
Lima	239	200	90	22	-149	18	**0**
Loreto	-42	84	4	-129	-47	52	**1**
Madre de Dios	0	-55	59	3	62	39	**-1**
Moquegua	-6	61	-118	9	-6	34	**5**
Pasco	157	151	309	200	20	-224	**1**
Piura	41	37	80	78	-30	-67	**1**
Puno	61	127	-17	151	1	-1	**1**
San Martín	-4	-10	-19	-29	4	43	**4**
Tacna	14	99	-186	-11	47	-10	**-1**
Tumbes	-75	65	-56	-17	13	181	**1**
Ucayali	-2	-41	-63	-17	4	-14	**-4**
National (Peru	40	-39	5	43	16	35	**1**

To exemplify the previous statement, between 2007 and 2017, average income per capita annual growth of the households in the lowest decile in Peru was over five percentage points higher than the one of the households in the top decile. Meanwhile, between 2012 and 2017, the difference was just around two percentage points, and many of the other lower deciles actually had an income growth similar to the top decile (see Figure 16). This change may be even worse in some regions.

Figure 16. Growth incidence curves. Peru, 2012 & 2017.



* The estimations are based on ENAHO.

There are however two positive conclusions. First, the adult ratio still contributed to reduce inequality in most regions, and this demographic effect has been accompanied by the adult employment ratio, which also decreased inequality in many regions when using this time frame. Second, and of vital importance for policymakers, public monetary transfers also continued to help reduce inequality in most regions (17 out of 25). Just as economic activity slowed down, social policy became more relevant to help curb inequality.

Now, regarding the geographical division, Figures 17 and 18 present the decomposition between 2012 and 2017. As in the previous case, the scenario has also varied. Labor income is not a uniformly equalizing force anymore, but the difference with respect to the political region division is that here it only contributes negatively to LMA (which is probably capturing the negative effect in Callao). In the other regions, labor income still fosters equality, but in a much lower relative magnitude than before. The graph also allows to distinguish the importance of public monetary transfers in this period. Again, the highlands and the jungle were the most benefited by them, which is cheerful in terms of the targeting of these transfers.

Figure 17. Absolute contributions to gains in equality (Gini points). Geographical regions, 2012 - 2017.

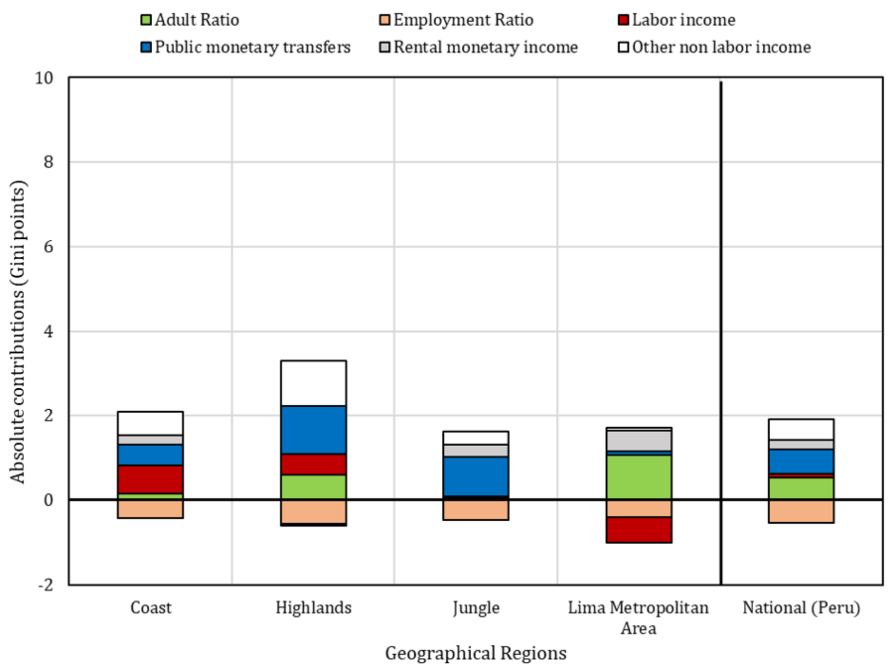


Figure 18. Relative contributions to gains in equality (%). Geographical regions, 2012 - 2017.

	Adult Ratio	Employment Ratio	Labor income	Public monetary transfers	Rental monetary income	Other non labor income	Gains in Equality (Gini points)
Coast	9	-25	39	30	13	34	2
Highlands	22	-21	19	41	-1	40	3
Jungle	4	-41	3	82	25	27	1
LMA	150	-56	-87	14	68	11	1
Peru	40	-39	5	43	16	35	1

6 Final remarks

This paper sheds light upon the regional dimension of inequality in Peru, both at the political and geographical level. This has been accomplished with three complementary analysis: (i) the description of trends of inequality at the regional level, (ii) the within-between decomposition of regional inequality, and (iii) the inequality decomposition into contributions of socioeconomic and demographic factors.

The initial descriptive part revealed that Peru has experienced indeed an important decrease in inequality between 2007 and 2017, but that the decline has not been steady. In fact, two important time periods were distinguished: the 2007-2012 marked by sharp reductions in inequality, and the 2012-2017, in which inequality barely changed. This evolution is supported in the literature. When looking at the trends within each region, the same story repeated itself: the last five years of the analysis implied a far smaller reduction in inequality.

The regional analysis also evidenced the large heterogeneities in inequality evolution across both political and geographical regions. Regarding the first ones, further research is needed to understand the causal mechanisms behind the heterogeneity. Meanwhile, with respect to the later ones, the analysis showed that the two richest regions (the coast and the Lima Metropolitan Area) were also the ones with higher equality. This means that any measure of well-being will see the coast and LMA ranking before the highlands and the jungle.

The second analysis, which corresponds to one of the main questions proposed at the beginning of the paper, showed a worrisome feature: the between contribution of inequality has risen steadily since 2011 for any measure of the Theil index and for any of the two regional classifications. This implies that as the inequality reduction within regions decreased, the importance of income divergence between regions became more important to explain the aggregate phenomenon. Policymakers looking to curb inequality should pay attention to this feature, because it means that some political regions are being increasingly left behind in the income distribution as economic and productivity growth in Peru is slowing down.

Meanwhile, the last quantitative analysis exposed the relative importance of socioeconomical and demographic factors in the inequality narrative at the national and regional level. When looking at the gains in equality between 2007 and 2017, it was possible to detect the importance of the demographic boom (fraction of adults in the households) and income growth (labor and private transfers) in curbing inequality. All regions benefited from the increase in the share of adults, and most of them saw a relative rapid increase in private income at the lower end of the income distribution. However, when the window of analysis was reduced to the last five years of the time frame (2012-2017), labor income contributions turned weaker, and in some regions it even helped increase inequality.

The last result suggests the importance of promoting productive jobs for the lower income households, and this is especially true when considering the fraction of employed adults. The employment ratio has not been a unanimous equalizing force across regions (both politically and geographically), meaning that in the 10-year and 5-year time frame, many of the poorest households saw an increase in their employment ratio not as high as in the richer ones. Thus, policymakers preoccupied by the recent slowdown in inequality reduction should consider the importance of creating productive jobs for the poorest individuals.

Finally, the decomposition exercise proved the importance of social policies in inequality reduction. In particular, public monetary transfers have helped curb inequality considering any time frame or regional classification in most regions and at the national level. In fact, during the last five years of the analysis, it was the second most important equalizing factor taking Peru as a whole. This hints at the good targeting of social policies, given that the highest contributions appear in the poorest regions, both politically (e.g. Apurímac, Ayacucho, Cajamarca and Puno) and geographically (the highlands and the jungle). Thus, these policies are having a strong redistributive effect that could be positive for long-term development.

References

Alarco, G., Castillo, C., and Leiva, F. (2019). *Riqueza y desigualdad en el Perú, Visión Panorámica*. Oxfam.

Araar, A. and Duclos, J.-Y. (2013). User Manual for Stata Package DASP: Version 2.3. Technical report.

Azevedo, J. P., Inchauste, G., and Sanfelice, V. (2013). Decomposing the recent inequality decline in Latin America. Policy Research Working Paper Series 6715, The World Bank.

Azevedo, J. P., Nguyen, M. C., and Sanfelice, V. (2012). ADECOMP: Stata module to estimate Shapley Decomposition by Components of a Welfare Measure. Statistical Software Components, Boston College Department of Economics.

Bourguignon, F. (1979). Decomposable Income Inequality Measures. *Econometrica*, 47(4):901–920.

Castillo, L. E. and Florián, D. (2019). Measuring the output gap, potential output growth and natural interest rate from a semi-structural dynamic model for Peru. Working Papers 2019-012, Banco Central de Reserva del Perú.

Chowell, G., Munayco, C. V., Escalante, A. A., and McKenzie, F. E. (2009). The spatial and temporal patterns of falciparum and vivax malaria in Perú: 1994–2006. *Malaria Journal*, 8(142).

Cruz Saco, M. A., Seminario, B., and Campos, C. (2018). Desigualdad (Re)considerada Peru 1997-2015. *Journal of Economics, Finance and International Business*, 2(1):13–52.

de Barros, R. P., de Carvalho, M., Franco, S., and Mendonça, R. (2006). Uma Análise das Principais Causas da Queda Recente na Desigualdade de Renda Brasileira. Discussion Papers 1203, Instituto de Pesquisa Econômica Aplicada - IPEA.

Escobal, J. and Ponce, C. (2012). Polarización y segregación en la distribución del ingreso en el Perú: trayectorias desiguales. Technical report.

Fields, G. S. (2001). *Distribution and Development, A New Look at the Developing World*. Russel Sage Foundation, New York, and The MIT Press, Cambridge, Massachusetts.

Gonzales de Olarte, E. (2010). Descentralización, divergencia y desarrollo regional en el Perú del 2010. In Rodríguez, J. and Tello, M., editors, *Opciones de política económica en el Perú 2011-2015*, Capítulos de Libros PUCP / Chapters of PUCP books, chapter 6, pages 175–204. Fondo Editorial - Pontificia Universidad Católica del Perú.

Herrera, J. (2017). Poverty and economic inequalities in peru during the boom in growth: 2004-14. In G. Carbonnier, H. Campodónico, S. T. V., editor, *Alternative Pathways to Sustainable Development: Lessons from Latin America*, pages 138–173. Brill | Nijhoff, Leiden, The Netherlands.

Instituto Nacional de Estadística e Informática (INEI) (2018). Evolución de la Pobreza Monetaria, 2007 - 2017. Informe técnico, Instituto Nacional de Estadística e Informática (INEI).

Jaramillo, M. and Saavedra, J. (2010). Inequality in post-structural reform peru: The role of market forces and public policy. In L. Lopez-Calva, N. L., editor, *Declining Inequality in Latin America. A Decade of Progress?*, pages 218–243. Brookings Institute Press, Washington, D.C.

Jaramillo, M. and Saavedra, J. (2011). Menos desiguales: la distribución del ingreso luego de las reformas estructurales. Technical report.

Mendoza, W., Leyva, J., and Flor, J. L. (2011). La distribución del ingreso en el perú: 1980 - 2010. In León, J. and Iguíñiz, J., editors, *Desigualdad distributiva en el Perú: Dimensiones*, pages 57–112. Fondo Editorial, Pontificia Universidad Católica del Perú, Lima, Perú.

Seminario, B., Zegarra, M. A., and Palomino, L. (2019). Estimación del PIB Departamental y Análisis de la Desigualdad Regional en el Perú: 1795 - 2017. Banco interamericano de desarrollo working papers, n° idb-wp-1016, Banco Interamericano de Desarrollo.

Yamada, G. and Castro, J. F. (2006). Poverty, inequality, and social policies in Peru: As poor as it gets. Working Papers 07-06, Centro de Investigación, Universidad del Pacífico.

Yamada, G., Castro, J. F., and Oviedo, N. (2016). Revisitando el coeficiente de Gini en el Perú: El rol de las políticas públicas en la evolución de la desigualdad. Working Papers 16-06, Centro de Investigación, Universidad del Pacífico.

7 Appendix

7.1 Gini estimates per year (%). Peru, 2007 -2017.

Year	Gini Index	[95% Conf.	Interval]
2007	50,0	48,5	51,4
2008	47,6	46,5	48,8
2009	47,3	46,1	48,4
2010	45,7	44,5	46,9
2011	44,9	43,9	45,9
2012	44,7	43,7	45,6
2013	44,1	43,3	44,9
2014	43,6	42,7	44,4
2015	43,5	42,6	44,4
2016	43,7	42,8	44,5
2017	43,3	42,5	44,2

7.2 Gini estimates per year and per political regions (%). Peru, 2007 -2017.

Point Estimates

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Dif.
Amazonas	49,5	48,5	50,7	49,9	45,9	46,6	48,5	46,0	46,1	45,9	43,2	-6
Áncash	45,9	45,7	41,8	45,3	45,1	42,7	43,2	42,0	40,1	41,5	42,0	-4
Apurímac	48,1	46,5	44,3	43,5	45,6	42,8	41,5	41,3	42,7	38,9	38,9	-9
Arequipa	44,6	43,2	42,9	42,1	39,4	40,5	36,5	37,9	36,3	38,2	34,8	-10
Ayacucho	50,3	46,6	46,6	45,7	51,1	46,2	46,3	45,2	46,1	45,0	44,1	-6
Cajamarca	52,0	53,7	51,2	53,0	53,1	52,2	52,3	48,5	49,9	48,2	48,4	-4
Callao	37,4	37,9	40,9	35,6	34,9	36,4	34,3	34,0	34,0	32,5	34,0	-3
Cusco	50,5	49,2	48,7	47,2	45,4	42,4	42,6	45,4	41,2	44,2	44,1	-6
Huancavelica	53,1	51,5	51,6	50,3	48,5	48,9	47,6	43,7	40,8	41,2	39,0	-14
Huánuco	51,7	50,4	47,0	49,3	51,2	52,5	50,6	46,9	47,6	46,9	46,9	-5
Ica	33,5	33,2	33,8	31,8	32,1	29,2	27,5	28,0	29,8	27,2	26,5	-7
Junín	46,0	49,4	41,8	39,6	38,7	40,1	38,8	39,2	40,4	44,2	41,6	-4
La Libertad	54,0	46,2	46,8	43,6	42,7	45,4	44,5	42,2	44,0	43,2	41,5	-12
Lambayeque	45,1	42,1	39,0	38,9	38,4	39,3	38,2	38,5	39,4	37,4	38,2	-7
Lima	46,3	43,0	44,3	42,9	41,9	41,2	40,8	40,5	40,5	41,0	40,6	-6
Loreto	51,5	50,4	54,7	52,6	50,5	52,1	48,8	47,7	49,8	50,1	51,4	0
Madre de Dios	39,5	40,9	39,2	37,0	41,8	39,8	38,7	39,7	39,8	35,2	40,7	1
Moquegua	51,2	51,4	49,3	48,3	47,4	46,0	43,9	45,4	43,6	42,3	41,6	-10
Pasco	48,9	51,0	47,9	42,5	45,1	41,0	45,1	40,3	39,3	38,4	40,4	-9
Piura	46,5	43,6	42,9	43,3	42,8	42,7	41,3	40,3	38,3	39,2	41,0	-5
Puno	46,4	44,6	44,8	42,5	41,5	43,2	46,1	42,9	40,5	40,5	42,3	-4
San Martín	50,4	49,2	51,3	48,1	48,9	48,4	50,5	51,0	49,1	46,0	44,5	-6
Tacna	41,6	43,1	41,8	38,9	39,7	39,4	40,1	40,4	41,9	39,9	40,2	-1
Tumbes	38,1	34,8	34,8	34,7	35,7	34,7	35,7	34,3	34,7	32,8	34,1	-4
Ucayali	45,4	43,1	40,4	36,7	34,3	34,1	34,5	33,1	35,9	33,9	37,5	-8

Note. The last column is the difference between the Gini coefficient in 2017 and 2007.

Std. Errors

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Amazonas	1,8	1,7	1,8	3,4	1,0	1,1	1,0	1,4	1,4	1,3	1,2
Áncash	1,6	1,5	1,5	1,9	1,4	1,1	1,3	1,3	1,1	1,1	1,2
Apurímac	1,6	1,9	1,7	1,4	1,5	1,3	1,2	1,5	2,7	1,6	1,1
Arequipa	1,4	1,3	1,3	2,3	1,6	1,4	0,9	0,9	1,0	1,1	0,8
Ayacucho	1,8	1,6	1,5	1,1	1,8	1,4	1,5	1,4	2,6	1,3	1,3
Cajamarca	1,6	1,9	1,3	2,0	1,8	1,5	1,5	1,6	1,6	1,3	1,4
Callao	1,5	1,6	3,3	1,4	1,0	1,3	1,0	0,9	0,9	1,1	0,9
Cusco	1,2	1,8	1,9	1,6	1,6	1,3	1,2	1,4	1,3	1,2	1,1
Huancavelica	2,1	2,9	3,2	2,8	2,7	2,3	2,3	1,9	1,5	1,6	1,2
Huánuco	1,7	1,7	1,6	1,4	2,0	1,6	1,5	1,4	1,4	1,4	1,4
Ica	1,2	1,3	1,2	1,0	1,3	1,2	0,8	0,8	1,1	0,8	0,8
Junín	1,6	2,3	1,4	1,4	1,1	1,3	1,1	1,0	1,3	1,6	1,2
La Libertad	5,7	2,1	2,0	2,0	1,6	1,7	1,2	1,3	1,3	1,0	1,1
Lambayeque	1,6	1,5	1,7	1,9	1,4	1,3	1,2	1,1	1,3	1,0	1,2
Lima	1,3	1,3	1,2	1,3	1,0	1,0	0,8	0,8	0,9	0,8	0,9
Loreto	2,2	1,5	1,9	1,5	1,8	1,5	1,4	1,4	1,5	1,5	1,5
Madre de Dios	1,7	1,8	2,2	1,4	1,9	2,0	1,5	1,7	2,0	1,4	2,1
Moquegua	2,6	2,3	2,2	2,5	2,1	1,8	1,8	1,8	1,7	1,3	1,3
Pasco	1,8	2,4	2,2	1,6	1,3	1,3	1,5	1,2	1,2	1,0	1,2
Piura	1,5	1,4	1,6	1,5	1,5	1,6	1,2	1,6	1,1	1,1	1,1
Puno	1,2	1,1	1,6	1,5	2,0	1,5	1,7	1,5	1,2	1,1	1,5
San Martín	1,5	1,5	2,2	1,3	2,0	1,6	1,9	3,4	1,4	1,3	1,3
Tacna	1,7	1,9	2,0	1,4	1,4	1,4	1,6	1,4	1,6	1,0	1,4
Tumbes	1,7	1,1	1,3	2,1	1,2	1,2	1,3	1,1	1,2	1,1	0,9
Ucayali	1,5	1,7	1,6	1,4	1,3	1,3	1,3	1,1	1,3	1,2	1,3

7.3 Mean real gross household income per capita by region. Peru, 2007 -2017. Soles of 2017.

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	# in 2017
Amazonas	482	488	570	618	601	620	606	597	656	678	688	17
Áncash	657	663	701	818	810	823	858	866	813	841	801	13
Apurímac	344	373	376	464	464	494	571	609	643	613	602	23
Arequipa	1002	1088	1097	1149	1180	1207	1231	1235	1200	1197	1200	3
Ayacucho	427	462	503	567	579	557	590	587	648	640	616	22
Cajamarca	390	450	484	536	568	591	577	566	579	567	594	24
Callao	849	886	1009	973	927	968	1020	1034	1034	1074	1050	4
Cusco	586	600	685	680	745	842	846	846	792	823	758	15
Huancavelica	282	315	363	430	478	488	508	441	476	498	477	25
Huánuco	456	512	505	566	606	690	710	690	699	689	685	18
Ica	778	814	892	919	956	938	946	978	972	998	983	7
Junín	658	748	710	723	827	824	839	832	870	904	832	11
La Libertad	897	782	864	859	834	897	924	951	952	982	1013	6
Lambayeque	693	702	716	732	748	786	786	813	866	934	905	10
Lima	1068	1072	1112	1126	1143	1219	1213	1239	1259	1312	1275	1
Loreto	480	506	521	588	606	638	627	629	620	620	660	20
Madre de Dios	775	823	874	944	1086	1188	1239	1140	1028	935	944	9
Moquegua	1019	1070	1099	1278	1286	1477	1409	1412	1366	1307	1216	2
Pasco	504	555	636	690	703	640	625	632	631	712	704	16
Piura	638	647	739	743	777	804	765	775	748	758	793	14
Puno	460	501	541	580	607	665	726	708	635	648	632	21
San Martín	563	618	622	717	759	779	777	760	746	787	831	12
Tacna	904	1005	969	1043	989	1067	1079	1039	1009	1004	1021	5
Tumbes	984	800	839	910	980	1000	975	940	945	1005	980	8
Ucayali	585	588	640	624	684	681	681	660	695	694	676	19

Note. The last column is the ranking in income per capita in 2017

7.4 Gini estimates per year and per geographical regions (%). Peru, 2007 -2017.

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Point Estimate											
Coast	44,0	39,8	40,2	39,7	38,1	38,5	38,0	37,1	37,5	36,7	36,7
Highlands	52,5	52,3	49,7	48,5	48,7	47,9	47,3	46,4	45,2	45,7	45,2
Jungle	48,8	48,0	48,8	46,4	46,3	46,4	46,8	45,3	46,0	45,4	45,3
Lima Metropolitan Area	45,6	42,9	44,0	42,5	41,6	41,3	40,5	40,2	40,4	40,6	40,4
Std. Error											
Coast	2,1	0,7	0,8	0,9	0,6	0,7	0,5	0,6	0,6	0,4	0,5
Highlands	0,6	0,7	0,6	0,6	0,6	0,5	0,5	0,5	0,5	0,5	0,4
Jungle	0,9	0,8	0,9	0,7	0,9	0,7	0,8	1,0	0,7	0,7	0,7
Lima Metropolitan Area	1,2	1,3	1,2	1,2	1,0	1,0	0,8	0,8	0,9	0,8	0,9

7.5 Relative contribution to inequality by political region (%). Theil's L index. Peru, 2007 - 2017.

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Amazonas	1,4	1,5	1,7	1,7	1,4	1,5	1,7	1,5	1,5	1,5	1,3
Áncash	3,2	3,5	2,8	3,5	3,7	3,4	3,5	3,4	3,1	3,2	3,4
Apurímac	1,4	1,4	1,3	1,3	1,5	1,3	1,2	1,3	1,4	1,1	1,1
Arequipa	3,2	3,3	3,4	3,5	3,1	3,3	2,8	3,1	2,7	3,1	2,5
Ayacucho	2,2	2,0	2,0	2,1	2,9	2,3	2,4	2,3	2,5	2,3	2,2
Cajamarca	5,3	6,3	5,7	6,7	7,0	6,7	6,8	6,0	6,3	5,7	5,8
Callao	1,7	1,9	2,3	1,9	1,8	2,1	1,8	1,9	1,9	1,7	1,9
Cusco	4,4	4,5	4,4	4,5	4,3	3,9	4,0	4,6	3,8	4,2	4,2
Huancavelica	1,8	1,8	1,8	1,9	1,8	1,8	1,8	1,5	1,3	1,3	1,2
Huánuco	2,9	3,1	2,6	3,2	3,6	3,8	3,6	3,1	3,2	3,0	3,1
Ica	1,0	1,1	1,2	1,1	1,2	1,0	0,9	0,9	1,1	0,9	0,9
Junín	3,7	4,8	3,5	3,3	3,1	3,5	3,3	3,5	3,7	4,4	4,0
La Libertad	7,0	5,7	5,9	5,5	5,5	6,3	6,2	5,6	6,3	6,0	5,7
Lambayeque	3,2	3,1	2,6	2,8	2,9	2,9	3,0	3,0	3,2	2,8	3,0
Lima	24,5	23,2	25,6	26,0	25,7	24,6	25,1	25,6	25,7	26,5	26,5
Loreto	3,4	3,7	4,5	4,3	4,1	4,4	4,0	3,9	4,4	4,3	4,8
Madre de Dios	0,2	0,3	0,3	0,3	0,3	0,3	0,3	0,4	0,4	0,3	0,4
Moquegua	0,6	0,7	0,6	0,6	0,7	0,6	0,6	0,7	0,6	0,6	0,5
Pasco	0,9	1,2	1,0	0,9	1,0	0,8	1,1	0,8	0,8	0,7	0,8
Piura	5,1	4,8	4,9	5,2	5,4	5,4	5,1	5,1	4,5	4,7	5,2
Puno	3,9	3,9	4,1	3,9	3,8	4,3	4,9	4,4	3,9	3,9	4,2
San Martín	2,6	2,8	3,1	3,0	3,1	3,1	3,5	3,7	3,4	3,0	2,8
Tacna	0,7	0,8	0,8	0,8	0,8	0,8	0,9	0,9	1,0	0,9	0,9
Tumbes	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4
Ucayali	1,3	1,3	1,2	1,0	0,9	0,9	1,0	0,9	1,1	1,0	1,2
Within	85,9	87,0	87,7	89,3	90,1	89,6	89,8	88,5	88,2	87,5	88,2
Between	14,1	13,0	12,3	10,7	9,9	10,4	10,2	11,5	11,8	12,5	11,8

7.6 Relative contribution to inequality by political region (%). Theil's T index. Peru, 2007 - 2017.

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Amazonas	0,9	0,9	1,2	1,4	1,0	1,0	1,0	1,0	1,1	1,1	0,9
Áncash	2,7	2,8	2,5	3,6	3,3	2,9	3,3	3,1	2,6	2,7	2,8
Apurímac	0,7	0,7	0,6	0,7	0,8	0,7	0,8	0,8	1,1	0,8	0,7
Arequipa	3,9	4,4	4,2	4,6	4,5	4,4	3,5	3,9	3,5	3,7	3,1
Ayacucho	1,3	1,2	1,3	1,4	2,1	1,4	1,5	1,5	1,9	1,5	1,4
Cajamarca	2,9	4,1	3,4	4,3	4,7	4,5	4,5	3,7	4,1	3,5	3,7
Callao	1,7	2,1	3,2	2,0	1,9	2,2	2,0	2,0	2,0	1,9	2,0
Cusco	3,1	3,5	3,6	3,5	3,6	3,3	3,4	4,1	3,0	3,5	3,2
Huancavelica	0,7	0,9	1,0	1,1	1,1	1,1	1,2	0,8	0,7	0,7	0,6
Huánuco	1,9	2,2	1,7	2,1	2,8	3,0	2,9	2,4	2,5	2,3	2,4
Ica	1,1	1,3	1,4	1,2	1,3	1,1	0,9	1,0	1,2	0,9	0,9
Junín	3,1	4,8	3,0	2,8	2,9	3,2	3,0	2,9	3,6	4,4	3,3
La Libertad	9,7	5,3	6,2	5,0	4,8	5,6	5,7	5,4	6,0	5,6	5,6
Lambayeque	2,7	2,7	2,3	2,4	2,4	2,5	2,5	2,5	2,9	2,7	3,0
Lima	37,4	35,9	36,8	37,6	35,6	35,8	36,2	36,1	37,1	38,4	38,5
Loreto	2,3	2,3	2,9	3,0	3,2	3,1	2,7	2,6	3,0	2,9	3,3
Madre de Dios	0,2	0,3	0,3	0,3	0,5	0,5	0,4	0,4	0,4	0,3	0,4
Moquegua	0,9	1,0	0,9	1,0	1,0	1,0	0,9	1,0	0,8	0,7	0,7
Pasco	0,6	0,8	0,8	0,6	0,7	0,5	0,6	0,5	0,5	0,5	0,6
Piura	4,0	3,9	4,2	4,5	4,7	4,6	4,0	4,5	3,5	3,6	4,2
Puno	2,2	2,4	2,6	2,6	2,9	3,0	4,1	3,3	2,5	2,5	2,8
San Martín	1,9	2,2	2,5	2,4	2,9	2,7	3,2	3,8	2,8	2,4	2,5
Tacna	0,9	1,2	1,1	0,9	1,0	1,0	1,1	1,1	1,1	1,0	1,1
Tumbes	0,5	0,4	0,4	0,4	0,5	0,5	0,5	0,4	0,5	0,4	0,5
Ucayali	0,9	0,9	0,8	0,7	0,7	0,7	0,7	0,6	0,8	0,7	0,8
Within	88,0	88,3	89,1	90,4	91,0	90,2	90,6	89,5	89,2	88,6	89,1
Between	12,0	11,7	10,9	9,6	9,0	9,8	9,4	10,5	10,8	11,4	10,9

7.7 Relative contribution to inequality by geographical region (%). Theil's L & T index. Peru, 2007 - 2017.

Region	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Theil's L Index											
Coast	17,1	15,2	16,1	16,7	16,3	16,5	16,7	16,5	16,7	15,9	16,3
Highlands	36,3	39,3	35,9	36,6	38,2	37,5	37,1	36,2	34,5	34,3	33,8
Jungle	11,8	13,0	13,9	13,2	13,4	13,8	14,4	13,7	14,4	13,8	14,0
Lima Metropolitan Area	23,7	23,2	25,1	25,6	25,3	24,8	24,9	25,4	25,6	26,3	26,4
Within	89,0	90,7	90,9	92,0	93,2	92,6	93,0	91,8	91,1	90,3	90,5
Between	11,0	9,3	9,1	8,0	6,8	7,4	7,0	8,2	8,9	9,7	9,5
Theil's T Index											
Coast	21,2	16,1	17,2	18,2	16,9	17,5	16,6	16,7	16,8	15,8	16,9
Highlands	24,4	29,7	27,0	27,7	30,1	28,7	29,2	28,5	26,8	26,4	25,3
Jungle	8,2	9,1	9,7	9,5	11,1	10,7	11,1	10,7	10,7	10,0	10,1
Lima Metropolitan Area	36,2	36,1	37,4	37,2	35,4	35,9	36,3	36,2	37,3	38,3	38,5
Within	90,0	91,1	91,3	92,6	93,5	92,8	93,2	92,1	91,5	90,7	90,9
Between	10,0	8,9	8,7	7,4	6,5	7,2	6,8	7,9	8,5	9,3	9,1