Economic complexity and Inequality: A spatial analysis for the Peruvian regions

Introduction

Recent empirical literature has risen in order to study the relation between a country's productive structure and its socioeconomic outcomes. Authors like Hidalgo & Hausmann (2009), Hartman et al (2016) and Acemoglu and Robinson (2012) link a country's level of economic complexity to socioeconomic development. More recently, Hidalgo (2016) analyzes, at country level, the relation between the Economic Complexity Index and Income Inequality. Showing that there is, indeed, an association between a country's productive structure and socioeconomic outcomes other than productivity levels and growth.

Although there have been other studies that conduct a similar analysis, most of them have concentrated in the country level neglecting the differences and heterogeneity that reside within a country regional level. Only Gao & Zhou (2017) and Morais, Stewart & Jordan (2020) analyze the regional case for China and Brazil. Their results show that the regional level of economic complexity is associated with lower Gini and Theil indexes. More specifically, they show what appears to be an inverted U relation between economic complexity and income inequality, suggesting the existence of a threshold of economic complexity in order to see the reduction of inequality.

Even when the two previously mentioned papers use subnational level observations for their analyses, neither of them takes into account the potential spatial spillovers that might take place between regions. Thus, this study analyses the case for Peruvian regions and uses a spatial econometric approach to consider the potential spatial interactions at the subnational level.

Theoretical framework

As stated by Auyang (1998) and Matutinovic (2010), the concept of complexity often refers to systems that are self-organized with multilevel structures and with the capability of changing abruptly to adapt to the external environment. The characteristics these authors identify for complex systems are circular causality, feedback loops and non-linear cause-effect response. This concept cannot be exempt from the interactions between agents within the system and, hence, the formation of networks is inevitable. According to Potts (2000) and Matutinović (2005) industrialized regions are complex systems that are not in equilibrium composed by a network of agents where nodes represent firms and the interactions within this system are business relations among firms. Consequently, these relations cannot be understood without the concept of markets, because they enable the functional connections of networks, decentralized removal, replacements and addition of new firms (Ulanowicz 1997; 2009).

These concepts are condensed by Hidalgo & Hausmann (2009) in a single measurement they denominated "Economic Complexity Index" which proxies the variety of knowledge types within a single country through the analysis of the variety of its export basket. Economic complexity can also measure the quality of institutions, diversity of markets and strength of networks within a country (or a region). Therefore, a link between economic complexity, inclusion and inequality can be established. As Hartman et al (2016) points out, the relation between the evolution of productive structures and inclusiveness has been previously studied by authors such as Harold Innis (1970), Engerman and Sokoloff (1997) and Acemoglu and Robinson (2012) through a historic analysis of early productive structures and the quality of institutions that were developed according the favorable or unfavorable conditions of the environment.

Hartman et al (2016) argue that institutions are dependent on the type of industry, workers, firms and networks developed. They compare the institutions created in productive environments such as Silicon Valley and the ones generated in mining regions, stating that the ones generated in a mining region would not work efficiently in Silicon Valley. These differences are also present within sectors but between different regions or countries. Although, it could be the case that the differences in institutions between different productive regions are higher that the difference between sectors. Using all these arguments, Hartman et al (2016) state that countries with more sophisticated industrial structures foster inclusivity and could present lower levels of income inequality.

Empirical literature

The literature studying the link between economic complexity and economic development is fairly recent and is growing rapidly. The variables that have been used most frequently are at country level and measure dimensions such as economic inequality, human development, and gross domestic product.

The first work that relates the variables of development and economic growth with the concept of economic complexity is that of Hausman et al (2013). These authors use a panel of data to carry out regressions on the GDP of various countries in the world (both developed and developing) and find that its explanatory variable, the “Economic Complexity Index” (ECI), is a fairly robust predictor for national per capita income. Within their research, they contrast this measure, developed by themselves, with variables such as the level of human capital and variables commonly used to explain economic growth at the country level. Their results suggest that the ECI is a much more robust predictor than the mainstream used variables.

Following a similar line, but using another measure of economic development, is Hartman et al (2016). These authors seek to answer the question of whether the productive structure determines (or if, at least, it is related to) the capacity of a country not only to generate but also to distribute income. They apply a panel data model using the GINI EHII base that has a fairly complete country-level sample of the Gini coefficient. The measure of economic complexity is, again, the ECI. Their results indicate that an increase in one standard deviation in economic complexity is associated with a reduction in Gini of 0.03. Which is equivalent to three more years of schooling. These results remain robust to the inclusion of income and human capital measures. Similarly, Le Caous & Huarng (2020) use the human development index for developing countries as a dependent variable. Their results indicate that economic complexity is positively correlated with human development and that this effect is channeled by the inverse relationship between complexity and income inequality.

Two papers that refine the estimations and results analysis using human capital, education levels and government spending are Lee & Vu (2019) and Chu & Hoang (2020). The first one, refine the estimates by including various measures of human capital. The results found by these authors show heterogeneous effects between developed and developing countries. The interaction between human capital measurements and the economic complexity index has a positive effect for developed countries and a negative effect for developing countries in the reduction of income inequality. The second paper identifies the thresholds of variables such as level of education, government spending and trade openness that facilitate the attenuating effects of economic complexity on income inequality.

A different line of recent empirical work focuses on business cycles and production volatility at the country level. The works that stand out in this body of empirical literature are those of Canh

& Tanh (2020) and Guneri & Yalta (2020). The first two authors use the economic cycles of 70 countries in the period from 1996 to 2014. The data panel used contains high-, middle- and low-income countries. The findings in this article suggest a negative relationship between economic complexity and the magnitude of business cycles. However, when analyzing by segmenting countries by income levels, they find that this relationship is significant only for high-income countries. Additionally, Guneri & Yalta (2020) used an auto-regressive vector methodology for a panel of 61 countries in the period 1986 to 2015 to analyze the relationship between GDP volatility and economic complexity. Similarly, these authors find a negative relationship between GDP volatility and the economic complexity index. Its results are robust to the use of various complexity measurements.

Finally, only two articles have been found that focus on outcome variables at the subnational level. First, Gao & Zhou (2017) use a data pool with 31 Chinese provinces for the period 2000 to 2015. Their results indicate that the economic complexity index explains almost 50 percent of the variance of the log of per capita GDP. The second study comes from Morais, Stewart & Jordan (2020). These authors use a panel of data with the 27 federative units of Brazil to estimate the relationship between the economic complexity index and income inequality (measured with the Gini index and Theil coefficient). This work points out the negative relationship between economic complexity and income inequality. However, this relationship is more accentuated for the federative units with relatively higher incomes.

Data and Model

Since the literature that analyses the relationship between economic complexity and income inequality has vastly focused on cross-country data (only Morais, Stewart & Jordaan, 2020) analyze the case of Brazilian states) the main objective of this paper is to analyze this relationship at the subnational level. Due to the fact that aggregated trade statistics could be misleading in representing the heterogeneity of a country's productive structure. The concentration of industries, agglomeration effects, and human capital externalities are seen through a more disaggregated lens (Moretti, 2004a, b; Winters, 2014; Figueroa, 2018; Del Carpio, 2020). Additionally, Balland et al (2020) finds that, at least for the case of the US, complex industries cluster in large cities. Thus, the analyzed effect at country level could be misleading because of this regional and subnational heterogeneity. This gives us a fair reason to suspect that the link between economic complexity (i.e., the productive structure) of a country and income inequality could also show some regional heterogeneity.

The regression model is based on the previous literature Hartmann et al (2017), Castilho et al. (2012) and Morais, Stewart and Jordan (2020). These papers give a good framework on the variables that should accounted for in the model and the interaction term between the measurements of human capital and the ECI are taken from Lee and Vu (2019). However, none of these authors used a spatial approach to observe the potential interactions between regions that might affect the relationship of economic complexity and income inequality.

$$y_{it} = \alpha_i + \beta_1 ECI_{it} + \beta_2 ECI_{it}^2 + \beta_3 Work.Univ_{it} + \beta_3 Work.Tech_{it} + \beta_4 Rural_{it} + \beta_5 Lnpop_{it} + \sum_{j=6}^{10} \beta_j \, GDP_{it}^j + \varepsilon_{it} \tag{1}$$

Table DM. 1. Variable definition and sources

Variable	Definition	Source
Gini	Gini coefficient	Calculated from the Peruvian National Household Survey (ENAHO)
Theil	Theil coefficient	
ECI	Economic complexity index	Calculated from the Peruvian export data (PROMPERU) and world trade data (COMTRADE)
Work Univ.	Participation of workers with some degree of university education	Computed from the Peruvian National Household Survey (ENAHO)
Work Tech.	Participation of workers with some degree of technical education	
Rural	Proportion of rural habitants	
Population	Population in millions	
GDP commerce	Per capita GDP of commerce	Calculated using data from the Peruvian National Institute of Statistics and Informatics (INEI)
GDP agriculture	Per capita GDP of agriculture	
GDP mining	Per capita GDP of mining	
GDP manufacture	Per capita GDP of manufacture	
GDP government	Per capita GDP of government	
GDP construction	Per capita GDP of construction	

The dependent variables to measure income inequality are the Gini and Theil indexes. Since both have some issues with sensitivity to transfers and variations in the tails of the distribution (Morais, Stewart & Jordaan, 2020) we decide to use both measurements in order to assure some robustness in our results. The ECI is calculated using a twofold criterion, the revealed comparative advantage and if the total value of export of a given product is higher than fifty thousand American dollars. Then the product complexity indexes (calculated from the COMTRADE datasets) are assigned to the ones that comply with at least one of the two criteria previously mentioned. The regional trade data is extracted from the Promotion of Exports division of the Ministry of International Trade and Tourism (PROMPERU) which has district level export data.

Tables DM 2 and 3 contain the means of the variables used in the model at the regional and year level. Following the revised literature, some measurements of human capital are included in the model. These are the labor force participation of workers that have some degree of university or technical formation. To account for agglomeration and urbanization effects we introduce the natural logarithm of the population and the proportion of the population that is on the rural size of a given region. Finally, the per capita gross products of commerce, agriculture, mining, manufacture, government and construction are also included in the model.

Both measurements of income inequality seem to have a relatively small volatility between regions through time. The Ica region has the lowest averages for both inequality measurements and, contrastingly, Cajamarca has the highest indexes of inequality of the twenty-four compared regions. The economic complexity index has, also, a relatively small variation between regions with Lima, Cusco and Loreto having the highest indexes of the group. Lima has the highest proportion of university and technical institute graduates and it's the most populated region of Peru. The typically mining regions have the highest per capita gross products (Madre de Dios, Pasco, Moquegua and Ancash).

Table DM 3 shows that income inequality has decreased modestly over time for all Peruvian regions. However, the productive structure appears to have remained the same in the span of time that is used for the analysis. The human capital variables have also improved in through time, the proportion of workers with some degree of higher education has doubled (from 15 to 30 percent in the last 15 years). Even though Peruvian population has increased steadily the proportion of rural habitants has decreased steadily in the analyzed period. The production over sectors has also increased over time.

Results

The first part of the discussion about the empirical findings will be using a panel approach taking into account the most commonly used variables in the empirical literature. Table X contains the first set of estimates that come from the initial model.

Table R1. Estimated coefficients primary regressions

	OLS		FE		RE	
Variables	Theil	Gini	Theil	Gini	Theil	Gini
ECI	0.028*	0.013	-0.000	-0.007	0.007	-0.001
	(0.015)	(0.012)	(0.013)	(0.009)	(0.012)	(0.009)
ECI_sqr	0.003	0.002	-0.009***	-0.006***	-0.006**	-0.005**
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Work Univ.	-0.003	-0.035	0.423*	0.330**	0.369**	0.309**
	(0.127)	(0.104)	(0.222)	(0.160)	(0.185)	(0.138)
Work Tech.	-1.861***	-1.300***	-0.888***	-0.476***	-0.950***	-0.491***
	(0.257)	(0.202)	(0.251)	(0.181)	(0.232)	(0.170)
ECI x Work Univ.	0.168*	0.096	0.082	0.046	0.081	0.032
	(0.089)	(0.074)	(0.114)	(0.083)	(0.107)	(0.079)
ECI x Work Tech.	-0.363**	-0.129	-0.389***	-0.188*	-0.386***	-0.193*
	(0.162)	(0.139)	(0.142)	(0.103)	(0.141)	(0.103)
LnPop	0.024***	0.016***	0.044	0.077*	0.028***	0.021***
	(0.005)	(0.004)	(0.059)	(0.043)	(0.008)	(0.006)
Rural	0.021	-0.012	0.017	-0.010	0.112**	0.087**
	(0.032)	(0.025)	(0.112)	(0.081)	(0.047)	(0.036)
GDP commerce	0.001	-0.011	0.008	0.007	0.011	0.004
	(0.010)	(0.008)	(0.013)	(0.009)	(0.012)	(0.009)
GDP agriculture	0.028***	0.014*	0.038***	0.018**	0.033***	0.016**
	(0.008)	(0.008)	(0.011)	(0.008)	(0.010)	(0.007)
GDP mining	0.003	0.003**	-0.003	-0.002	-0.002	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
GDP manufacture	0.015***	0.013***	-0.001	-0.001	0.001	0.000
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.002)
GDP government	0.033	-0.001	0.013	0.019	0.020	0.002
	(0.033)	(0.027)	(0.039)	(0.028)	(0.033)	(0.024)
GDP construction	-0.009	-0.013	-0.019*	-0.018**	-0.015	-0.016**
	(0.012)	(0.011)	(0.010)	(0.007)	(0.009)	(0.007)
R-squared	0.678	0.657	0.686	0.691	0.682	0.685

N=24, T=15; Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The first two columns of **Table R1** show the estimates coming from a pooled OLS estimation using the 24 Peruvian regions. These first regressions show that ECI is not significant (the first one even shows that ECI is positively associated with income inequality). Nonetheless, these first two models do not incorporate the potential effects that are proper to every region. Thus, a panel approach must be used in order to attempt controlling this specific effects.

The following four rows of **Table R1** show the estimates coming from the fixed effects panel and random effects panel. Since the comparison between models is not the focus of this paper, we will analyze the variables that stay robust through most of the estimations. The first important finding in these regressions shows that there is an inverted U form of the degree of diversification and income inequality. The linear term of the ECI is not significant in all regressions. However, the quadratic term of ECI is significant at the 0.01 level showing robustness as it remains significant in all of the panel regressions. Moreover, the size of the quadratic term is quite small. But this reduced size could be consistent with the productive structure of Peru. As we have shown before, the ECI has varied very little in the scope of time analyzed and the main part of Peruvian exports come from primary industries which have small values of complexity.

The measurements that attempt to capture the level of human capital also show some relatively small volatility between models. Similarly, to the empirical literature, there is a side of human capital (the concentration of workers with some degree of university education) that seems to deepen the observed income inequality in the Peruvian regions. Conversely, the participation of the workers with some degree of technical education seem to attenuate the magnitude of income inequality at the regional level. The two interactions terms used in these regressions differ on their importance to reduce income inequality. The only term that remains significant is the interaction between the participation of workers with some degree of technical education and the measurement of economic complexity. This would point out that regions with a greater participation of technical human capital the effect of the productive structure on income inequality becomes stronger and significant.

The production measurements seem to be quite relevant throughout the entire body of this first estimations. Their effects are consistent with one would expect. For example, the per capita gross product of agriculture seems to exacerbate income inequality at the regional level. On the contrary, construction seems to reduce income inequality. Since there are differences on the entanglements that both these industries have, their estimated effect on income inequality seems adequate.

Spatial regressions

Another relevant analysis to do with the data available is the incorporation of space into it. To perform this analysis, first we observe if there are a sign of spatial regularities. Figure 1 shows the evolution of the density for the GINI coefficient between 2004 to 2019 on the left side and the Theil index for the same period of time on the right side. The figure shows evidence that regional GINI and the Theil index suffered two noticeable changes. First, there is a small reduction in the national GINI and Theil average along these years; second, the GINI coefficient distribution changed from unimodal distribution to multimodal one, i.e., the GINI for a group of regions converge toward a certain level while another group of regions converge to a different GINI average coefficient. On the other hand, the Theil coefficient showed a less heterogenous behavior during the same period of time.

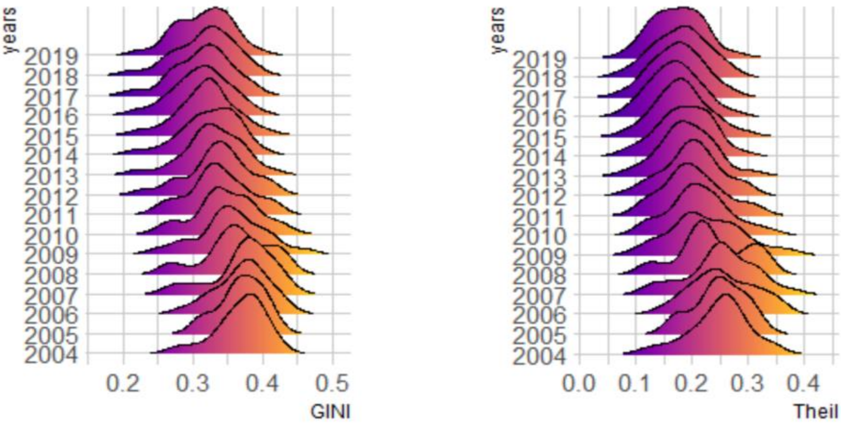
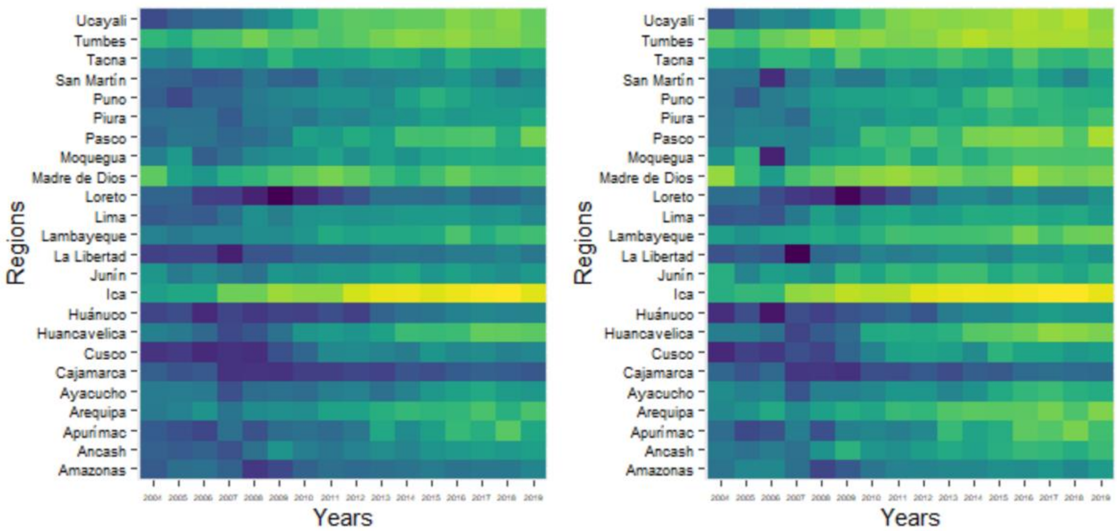


Figure 1. Evolution of the GINI and Theil coefficient distribution among regions: 2004-2019

The evidence presented in Figure 1 is supported by Figure 2 in which we show the evolution of the GINI coefficient and Theil index by regions along the period of analysis. In this figure we observe that regions such as Ica has low GINI and Theil index while other regions such as Amazonas, Cajamarca, or Loreto, the GINI and Theil indicators have higher values. In other words, there are some regions with low income inequality and regions with high income inequality. This disparity among the structure of income distribution across regions allow us to study the possibility that space has some effects on our analysis.



A first spatial statistical approach of the variables may be done by using the Moran's Indicator for each of the relevant variables on our analysis. Table 2 shows the results of estimation the Moran's indicator for the GINI coefficient, the Theil index, and the ECI. Also, we present the p-values corresponding to each estimated statistic. For each of the years, the table does not show evidence of spatial relationship between Peruvian regions for these three variables, in all the cases the p-value is not below 0.05. Nevertheless, this Moran's indicator is limited since it only considers the relationship across the space among regions in each period of time and ignores potential effects of time. To confirm these preliminar results we test the spatial behavior of our variables by using a local cross-sectional dependence test (CP test) and the randomized W test to confirm CP results.

Table 2. Moran's Indicator

year	gini	theil	eci
2004	-0.082	-0.141	-0.078
p-value	0.615	0.767	0.601
2005	0.012	-0.015	-0.043
p-value	0.339	0.416	0.499
2006	-0.157	-0.251	-0.149
p-value	0.804	0.940	0.785
2007	-0.088	-0.077	-0.241
p-value	0.631	0.600	0.931
2008	-0.051	-0.058	-0.218
p-value	0.522	0.544	0.905
2009	-0.050	-0.057	-0.069
p-value	0.518	0.540	0.575
2010	-0.059	-0.068	-0.115
p-value	0.545	0.572	0.705
2011	-0.022	-0.024	-0.146
p-value	0.436	0.440	0.781
2012	-0.067	-0.101	-0.090
p-value	0.569	0.667	0.637
2013	0.007	0.013	-0.080
p-value	0.352	0.335	0.610
2014	-0.003	-0.011	-0.239
p-value	0.382	0.404	0.929
2015	-0.034	-0.045	-0.070
p-value	0.472	0.503	0.580
2016	0.019	0.031	-0.113
p-value	0.318	0.287	0.700
2017	0.069	0.033	-0.088
p-value	0.200	0.282	0.632
2018	-0.006	-0.006	-0.193
p-value	0.389	0.390	0.870
2019	-0.002	0.000	-0.245
p-value	0.377	0.373	0.935

In Table 3, we observe the results for CP and RW test to test the possibility of having spatial effects across regions. The main characteristic and strength of these tests is the capability to incorporate the structure of panel data in it. In fact, in the CP test, the null hypothesis is no cross-sectional dependence against the alternative hypothesis of local cross-sectional dependence; in other words, we are contrasting whether there is no spatial interaction among regions (null hypothesis) versus spatial interaction (alternative hypothesis). However, this test is sensitive to serial correlation or non-spatial types of dependence. To confirm the results obtained with the CP test, we use the RW test, which bases its analysis on the idea that spatial correlation does not accept permutation among regions; otherwise, the correlation detected has non-spatial bases in it. In other words, the RW test has the null hypothesis of no spatial dependence versus the alternative hypothesis of spatial dependence.

Table 3. CP test results

	CP statistic	p-value	RW test (p-value)
gini	19.176	0.000	0.57
theil	19.135	0.000	0.494
eci	6.245	0.000	0.126
vab94_comercio_pc	15.985	0.000	0.356
vab94_agricultura_pc	0.013	0.990	0.124
vab94_mineria_pc	18.403	0.000	0.024
vab94_manufactura_pc	0.023	0.982	0.364
vab94_gobierno_pc	13.262	0.000	0.222
vab94_construccion_pc	20.437	0.000	0.954
vab_total_94_pc	24.277	0.000	0.532

On Table 3 we observe that under the CP test, the null hypothesis of no spatial cross-sectional dependence is rejected for most of the variables, i.e., there is evidence of local spatial correlation among regions. On the other hand, we do not reject the null hypothesis of no spatial dependence with the RW tests. In other words, the RW test is suggesting that the regularities we may found evidence by using the CP test are non-spatial ones.

This initial analysis is suggesting that the variables we are including in our analysis are not spatially correlated; however, there is some evidence of a potential non-spatial interaction among regions that must be investigated in more detail.

Table R2. Spatial lag panel (Theil)

Variable	Fixed effects with year dummies				Random effects without year dummies			
	Coeff.	SE	t	p-value	Coeff.	SE	t	p-value
Intercept					0.341	0.058	5.884	0.000
Lambda	-0.194	0.074	-2.613	0.009	0.160	0.042	3.851	0.000
ECI	0.001	0.012	0.096	0.923	0.008	0.013	0.603	0.547
ECI_sqr	-0.008	0.003	-2.757	0.006	-0.006	0.003	-1.957	0.050
Work Univ.	0.375	0.202	1.854	0.064	0.171	0.196	0.873	0.383
Work Tech.	-0.842	0.229	-3.676	0.000	-1.139	0.239	-4.761	0.000
ECI x Work Univ.	0.077	0.104	0.735	0.462	0.212	0.112	1.891	0.059
ECI x Work Tech.	-0.376	0.130	-2.891	0.004	-0.506	0.145	-3.499	0.000

LnPop	-0.018	0.103	-0.175	0.861	0.181	0.056	3.261	0.001
Rural	0.050	0.054	0.919	0.358	0.004	0.011	0.386	0.700
GDP commerce	0.007	0.012	0.593	0.553	0.035	0.010	3.583	0.000
GDP agriculture	0.037	0.010	3.587	0.000	0.029	0.011	2.681	0.007
GDP mining	-0.003	0.002	-1.681	0.093	-0.005	0.002	-2.141	0.032
GDP manufacture	-0.001	0.003	-0.188	0.851	0.002	0.004	0.560	0.576
GDP government	0.013	0.036	0.376	0.707	-0.045	0.028	-1.614	0.107
GDP construction	-0.020	0.009	-2.109	0.035	-0.042	0.009	-4.515	0.000

N=24, T=15; The Lambda shows the spatial lag coefficient of the endogenous variable

The spatial econometric analysis takes into account the spatial lag of the dependent variable and considers the interactions between neighboring regions and the effects on the overall sample of this interactions. Given that in both models the spatial lag coefficient is significant at the 0.01 level, we can say that there is indeed a spatial relation between the inequality of a given region and its neighbors. The coefficients of the linear term of the Economic Complexity Index remains not significant and the quadratic term reduces a bit. Nonetheless, even when controlling for spatial interactions on the dependent variable, the inverted U shape remains. Thus, going according to the claim of the empirical literature that there seems to be a threshold of complexity in the productive structure to see the its beneficial effects.

The human capital variables hold their pattern of significance between models and between measurements. Although the interaction term of university level human capital and the ECI becomes significant in the fixed effects model with spatial lag. The technical level human capital variable remains robust and negative and so does its interaction term with the ECI. Additionally, the other population measurements become non-significant with the exception of the logarithm of the population on the RE without year dummies model. The per capita GDPs hold a very similar structure in this model than on the previous one. The commerce GDP becomes negatively associated with the ECI and significant in the RE model.

Concluding remarks

The literature relating the productive structure and income inequality is quite recent, therefore a little scarce. However, there are some studies that mention this relationship at the country level for developed countries. Not surprisingly, the literature relating these two measures in developing countries is less available. The only study that analyses these two variables comes from Brazil where the authors use a subnational level of aggregation to account for heterogeneities within a country. Their findings suggest the existence of a threshold of economic complexity to see a reduction of inequality.

Given the little available literature and the lack of a more spatial based approach, this paper has conducted a spatial analysis of the regional data from Peru. Using the export data from the Peruvian IRS, the National Household survey and data from the Statistics Institute, a spatial model relating the Economic Complexity Index and the Gini and Theil indexes was constructed. Our results show an inverted U relation similar to that found in Brazil. The spatial lagged model does not impact significantly the main relation of the model but adds a bit more accurate effects due to the absorption of potential spatial effects between the Peruvian regions.

Given these results, policy recommendations go in line with the decentralization of production at the regional level and diversification of the productive structure utilizing the paths on the product space. This could mean a quicker arriving at the complexity threshold in order to perceive the reduction of inequality.

References

Hartmann, D., Guevara, M., Jara-Figueroa, C., Aristarán, M., Hidalgo C. (2016) Linking Economic Complexity, Institutions and Income Inequality. MIT Media Lab.

Matutinovic, I. (2010) Economic complexity and the role of markets. Journal of Economic Issues. XLIV (1)

Hidalgo C., Hausman, R. (2009) The building blocks of economic complexity. CID working paper series.

Bishop, A. & Mateos-Garcia, J. (2019) Exploring the link between economic complexity and emergent economic activities. National Institute Economic Review (249).

Balland, P., Jara-Figueroa, C., Petralia, S., Steijn, M., Rigby, D., and Hidalgo, C. (2020) Complex economic activities concentrate in large cities. Nature, Human Behaviour.

Hidalgo, C. (2021) Economic complexity theory and applications. Nature Reviews

Khanh-Chu, L & Phuong-Hoang, D. (2020) How does economic complexity influence income inequality? New evidence from international data. Economic Analysis and Policy (68): 44–57

Le Caous, E and Huarng, F. (2020) Economic Complexity and the Mediating Effects of Income Inequality: Reaching Sustainable Development in Developing Countries. Sustainability 12, 2089

Phuc-Canh, N & Dinh-Thanh, S (2020) The Dynamics of Export Diversification, Economic Complexity and Economic Growth Cycles: Global Evidence. Foreign Trade Review 1–27

Dogan, B., Driha, M., Balsalobre-Lorente, B., Shahzad. (2020) The mitigating effects of economic complexity and renewable energy on carbon emissions in developed countries. Sustainable Development. 1–12

Guneri, B. & Yalta, A. (2020) Does economic complexity reduce output volatility in developing countries? Bull Econ Res.1–21.

Bandeira-Morais, M., Swart, J. & Jordaan, J (2021) Economic Complexity and Inequality: Does Regional Productive Structure Affect Income Inequality in Brazilian States? Sustainability 13, 1006.

Hartmann, D., Guevara, M., Jara-Figueroa, C., Aristarán, M & Hidalgo, C. (2016) Linking Economic Complexity, Institutions and Income Inequality. MIT media Lab.