Regional Income Disparities, Distributional Convergence, and Spatial Effects: Evidence from Indonesia

Gunawan, Anang and Mendez, Carlos and Santos-Marquez, Felipe

24 November 2019

Noname manuscript No.
(will be inserted by the editor)

Regional Income Disparities, Distributional Convergence, and Spatial Effects:

Evidence from Indonesian Regions 2010-2017

the date of receipt and acceptance should be inserted later

Abstract Using a novel dataset constructed for this study, the spatio temporal dynamics of income per capita across 34 provinces and 514 districts in Indonesia over the 2010-2017 period are analyzed. First, an exploratory spatial analysis suggests that spatial autocorrelation is only significant at the district level, and it appears to be robust from 2013 to 2017. Consequently, at this level, a spatial filtering model is used for decomposing income into a spatially independent component and a spatial residual. Next, through the lens of a distributional convergence framework, it was found that the spatially independent component shows less intra-mobility than the original income variable. When analyzing beta and sigma convergence, strong converging patterns are found for filtered variables and the speed is higher for the filtered data. Thus, it is argued that neighbor effects have played a significant role in slowing the pace of income convergence at the district level. In terms of provinces, beta convergence is reported and the distributional convergence framework suggests the formation of at least three clubs with high intra-distribution mobility for the upper end of the distribution. The article is concluded by relating these findings to some plausible policy interventions.

Keywords Getis filter · Convergence · Distribution-based clustering · Nonparametric distribution · Indonesia
JEL Classifications O40 · O47 · R10 · R11

1 Introduction

The decentralized system has replaced centralized governance in Indonesia for about two decades. Through a more significant political power and authority in fiscal and local administration, a decentralized system was expected to improve public service delivery and economic performance of the regions. However, there is limited research about decentralization in Indonesia that has provided a conclusive

Address(es) of author(s) should be given

answer on the regional income disparity dynamics. Pepinsky and Wihardja (2011) find that extreme endowment heterogeneity, factor immobility, and lack of capacity on local government institutions undermine the high expectation of the successful decentralized government. However, on the other side, Talitha et al. (2019) argue that during the early years of implementation, the decentralization have a negative impact on regional development. The institutional reformation and some improvements during the process of decentralization bring positive outcomes, which is shown by reducing trend of inter-regional disparity and spatial fragmentation.

The discussion on regional income per capita disparity in Indonesia continues with an emphasis on whether regional per capita incomes tend to converge or diverge over time. Besides, the neighbour's role is becoming more important in a regional development context. Motivated by these issues, in this paper, we investigate the Gross Regional Domestic Product per capita (GRDP per capita) convergence process at the provincial and district levels during 2010-2017. Using a distributional convergence framework, spatial autocorrelation analysis, the Getis filter and the classical convergence framework, this study examines the regional dynamics of income and evaluates the spatial dependence across provinces and districts in Indonesia.

Based on the newly constructed dataset and our methodological approach, we aim to answer the following research question: to what extent the role of space affects the convergence dynamics across Indonesian regions? We refer to regions because we analyze spatial dependence at two different spatial scales, which are 34 provinces and 514 districts.

Our findings show that spatial autocorrelation is only significant at the district level. By comparing the dynamics of filtered and non-filtered data, this study also indicates that there is to some extent regional mobility, especially in the upper tails of the distribution. However, when considering the classical convergence framework, we recover both sigma and beta at the district level for the filtered variable, with a speed that is of similar magnitude as the one reported in the literature at the regional level. Moreover, it seems that spatial effects are responsible for slowing down an otherwise faster convergence process. In terms of provinces, the classical convergence framework indicates that disparities have been significantly reduced. Furthermore, the distributional dynamics perspective provides more information about 3 convergence clubs and the high mobility of the richest regions.

There are three contributions of this research. First, we utilize a novel database created as part of this project that includes all provinces and districts. This dataset also consists of a novel set of data points for the centroids of each district. Second, beyond studying the the income dynamics of the average region we analyze the intra-dynamics of the income distribution over time. Thirdly, we shed light onto the effects that neighboring regions have on the overall convergence process at the district level.

The rest of the paper is organized as follows. Section 2 and 3 present a literature review and the methods and data. Section 4 shows the results of the classical convergence perspective, the distributional convergence framework and the spatial filtering of the data. Lastly, section 5 includes some concluding remarks.

2 Literature Review

Recently, the convergence analysis has focused into the sub-national.For example, Rey and Gallo (2009) argue that the various forms of regional interdependence, such as capital and labor mobility along with trade flow, can be captured using a sub-national dataset and its one of its advantages. Some literature have extensively analyze the convergence process at the regional level, such as (Barro and Sala-i Martin 1991, 1992a; Rey and Montouri 1999; Fingleton 2004).

In Indonesian economy, large per-capita income differences across provinces is a well known issue (Esmara, 1975). Regional income differences seem to persist despite considerable policy efforts that aim to promote a more balanced growth path. Although a series of five-year development plans, fiscal transfer programs, internal migration policies, and integrated economic zones initiatives aimed to reduce regional disparities have been put in place; the outcomes still remain below their original targets (Akita 1988; Garcia and Soelistianingsih 1998; Kataoka 2012).

In the case of Indonesia, not many studies have been conducted that analyze the per capita income, especially at district level. Hendajany et al. (2017) , use Indonesian districts -level data over 277 districts in 25 provinces, analyse beta and sigma convergences in various period from 1986 to 2010. They found beta convergence in Indonesia using data for the 1986 to 2010 period. However, using the similar data set, the study shows that sigma convergence was not present during that time period in Indonesia. Another study with the district-level dataset, conducted by Aginta et al. (2020) to evaluate the prospect of regional income convergence in the post-decentralization era. The result indicates five clubs of convergence for 514 districts with the richest group dominated by big cities and districts rich in natural resources. Meanwhile, the poorest regions are mostly located in the eastern part and characterized by being island-based districts.

In terms of methodologies, the Getis filter has been used in similar studies on the economic development of regions and sub-regions in Europe such as Fischer and Stumpner (2010), in Brazil Cravo and Resende (2013) or in China Villaverde and Maza (2012); Yet, to the best of our knowledge, this is the first paper that studies the spatial filtering of income per capita for regions in Indonesia after the decentralization of the government.

There are some studies regarding spatial effects during decentralization period. One of the studies was conducted by Vidyattama (2013) to analyze unconditional beta convergence in order to capture the neighborhood effects Using a spatial autoregressive (SAR) lag model and a spatial error model (SEM). The study utilizes Real GDP per capita and the human development index dataset in 26 provinces and 294 districts. Through inverted distance matrix, the study found statistically insignificant neighborhood effect. Focusing on decentralization period, Aritenang (2014), uses spatial analysis to investigates the decentralization effect on regional development, focuses on economic convergence during 1994 to 2004. The study finds the presence of strong spatial autocorrelation among Indonesia districts. In addition, the study also reveals that the speed of regional convergence is higher during the decentralization period.

3 Data and Methods

3.1 Data

After implementing the decentralization system, the number of provinces and districts in Indonesia have changed rapidly. The number of Provinces increased from 26 to 34 Provinces during 1999-2017, while the number of districts (regencies and cities) increased from around 330 districts to around 514 districts. Indonesia Central Bureau of Statistics has published periodically Gross Regional Domestic Product (GRDP) data by regencies in Indonesia since 1998. However, the number of regencies published has changing during the period of publication. In order to obtain a balanced panel dataset, this study constructs Gross Domestic Product per capita using data from the Indonesia Database for Policy and Economic Research (INDODAPOER) World Bank.

We construct data for per capita GRPD from GRDP at constant price and population numbers by districts. The data has been updated from constant price 2000 to 2010 and the missing observation caused by creation of new districts are solved using interpolation. This study uses a linear regression method with year and reference provinces as regressors to do the interpolation. We assume that the new district is having co-movement with its reference district. The trend of the new districts refers to their actual data in the years when the data are available and follow its reference district in the year when they are interpolated. More details on the construction of the database can be found in the appendix.

Although the decentralization process started in the year 2000, we consider the data from 2010 to 2017, for several reasons. First, we aim to consider 3 5-year transition in the distributional convergence framework, similarly to the methodology used by (Fischer and Stumpner 2010) for European regions. However, there is one more constrain in the period 2000-2017 and it is related to the Global Financial crisis in 2008-2009. This crisis presented an structural break in the growth rate of income per capita in Indonesia and our aim is to study the decentralization policy and therefore we must remove the period during and right after the crisis as this external shock heavily affected income dynamics

Moreover, we could for example study the income convergence from 2000 to 2007. However, a robustness check of spatial autocorrelation shows that the differential Moran's I is negative in 2002-2001. Thus we prefer not to use data for those early years as the spatial dependence is structurally different when compared to the continuous positive value for both the differential and standard Moran's I in the 2010s (see figure 10). Lastly, as shown in Figure 16 the number of districts in 2000 was 341 while the number increased to about 500 by 2010 and further to 514 in 2017. The changing number of districts imply that a large number of data was interpolated for the years 2000 to 2010 and thus the reliability of this data for over 150 interpolated districts may not be as robust as for just 14 districts in the 2010s. Considering all these challenges, the most plausible avenue for a robust analysis of income dynamics seems to be restricted to the period 2010 to 2017.

The geographical structure of Indonesia is characterized by islands; in the case of provinces this study uses a distance parameter for the specification of neighboring

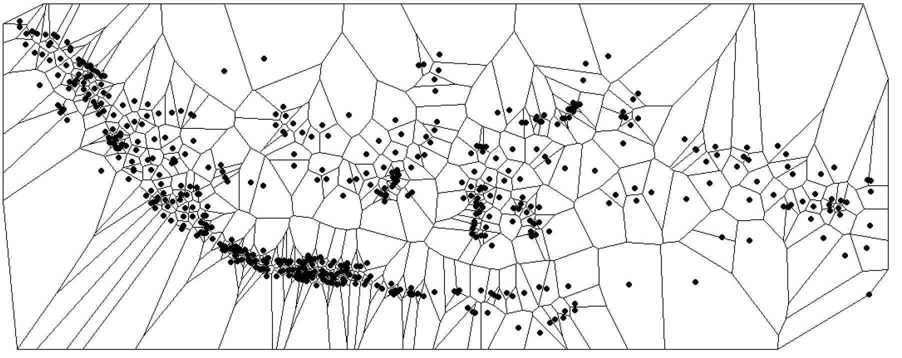


Fig. 1: Thiessen poligons for the centroids at district level

districts for calculating the spatial weights matrix and a queen contiguity criteria in the case of Districts. The Thiessen polygons needed to evaluate contiguity are recovered from the centroids of districts (see figure 1 The coordinates of the centroids are determined by the location of the capital cities of the districts based on the data from the Ministry of Internal Affairs and Geo-spatial Information Agency.

3.2 Classical Convergence Framework

The seminal convergence works of Barro (1991), Barro and Sala-i Martin (1991) and Barro and Sala-i Martin (1992a) started an ever increasing literature that aims to find convergence patterns across countries, regions, industries, etc, when analysing socio-economic variables. Although the original work by Barro and Sala-i-Matti was related to economic variables, it can be generalized as follows Barro and Sala-i Martin (1995):

$$(1/T) \cdot log\frac{y_{iT}}{y_{i0}} = \alpha - \frac{[1-e^{-\beta T}]}{T} \cdot log(y_{i0}) + w_{i,0T} \qquad (1)$$

Where i represents a region, 0 and T are the initial and final times, y is the studied variable, β is known as the speed of convergence, α includes unobserved parameters including the steady state and $w_{i,0T}$ is the error term. If there are robust signs of beta convergence according to (1), a different parameter known as the "half-life" can be evaluated:

$$half \cdot life = \frac{log2}{\beta} \qquad (2)$$

This parameter indicates the time that it would take the average region to reduce by half the distance between its initial state and the final equilibrium state.

A second measure of convergence is known as $\sigma - convergence$ (Quah 1993a), which takes into consideration the dispersion of the variables at the initial and final times. For analyzing $\sigma - convergence$ in this paper we use the standard deviation of the log of the RGDP per capita. Secondly, for robustness check we also use the gini coefficient to measure regional inequality. The beta and sigma convergence calculations are performed using the R package REAT (Wieland (2019)).

3.3 The Distributional Convergence framework

The study of the classical convergence hypothesis pioneered by Barro and Sala-i-Martin (Barro and Sala-i Martin (1991, 1992a); Barro (1991)) is useful for analysing the average long-run development of regions. However, in order to study the internal dynamics of the cross-section, a novel approach was suggested by Quah (1993b, 1996, 1997). This approach characterizes the dynamics of a system which is modeled using an stochastic kernel.

Figure 2 depicts a visual summary of the distributional dynamics framework. Utilizing stochastic kernels, the cross-sectional distribution and the internal dynamics of regions that are represented by a, b ,c and d in Figure 2 can be analysed. To formalize the framework shown in Figure 2, let us first define $Z_t(x)$ as the cross-regional GRDP per capita distribution at time t and $Z_{t+s}(y)$ as the distribution at time $t+s$. The internal dynamics of the distribution are computed by evaluating the first-order auto-regressive process of a time-homogeneous Markov chain shown in the following equation:

$$\underbrace{z_{t+s}(y)}_{Future\,Distribution} = \int \underbrace{z_{t+s|Z_t=x}(y)}_{Stochastic\,Kernel} \underbrace{z_t(x)}_{Initial\,Distribution} dx \tag{3}$$

The stochastic kernel estimation can usually be evaluated by dividing the joint probability density function $z_{t,t+s}(y,x)$ by the marginal probability density function $z_t(x)$:

$$z_{t+s|Z_t=x}(y) = \frac{z_{t,t+s}(y,x)}{z_t(x)} \tag{4}$$

One approach to calculate these parameters relies on the discretization of the Z space. However in this study we prefer to retain a continuous Z space and will analyze convergence and the presence of convergence clubs by plotting modal boxplots, as suggested by Hyndman et al. (1996). [1]. The plots related to this framework are performed using the R package HRDCDE (Hyndman (2019)).

3.4 Spatial Autocorrelation

In the academic literature several statistics are used for testing spatial autocorrelation. In this study the Moran's I statistic will be used. This statistic was suggested by

[1] For a more detailed presentation of the distributional approach see Magrini (2009); Fischer and Stumpner (2010); Mendez (2018, 2017)

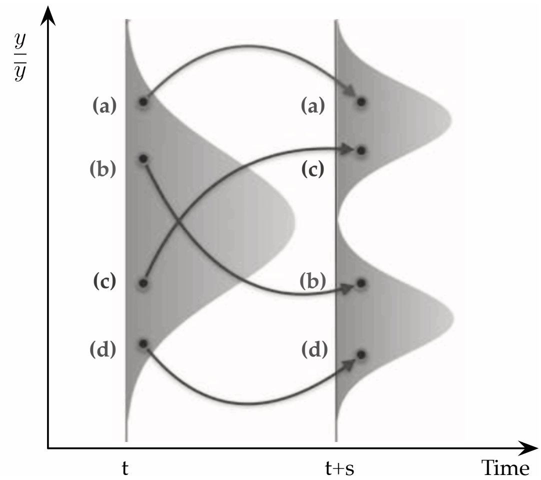


Fig. 2: Distribution dynamics framework: A graphical summary

Source: Adapted from Quah (1993).

Moran (1948) and can be defined as the coefficient of the correlation of the regression of the spatially lagged variable on the regular variable. Given a weight matrix $W = w_{ij}$ and a standardized variable z_i[2] which can be associated to each region i, the Moran's statistic I can be defined as:

$$I = \frac{\sum_i \sum_j w_{ij} z_i . z_j}{\sum_i z_i^2} = \frac{\sum_i (z_i \times \sum_j w_{ij} z_j)}{\sum_i z_i^2}$$

Given that the statistic is the coefficient of an OLS regression, it has an associated significance p-value. However given its spatial significance it also has a pseudo p-value obtained by a randomization simulation. For more details on this significance see the Geoda documentation Anselin et al. (2006).

In addition, one procedure to analyse the robustness of the standard Moran's I is to control for locational fixed effects. To remove these fixed effects the Moran's I for the variable $z_{i,t} z_{i,t1}$ is reported. This means that if at location i a fixed effect μ_i is present, it is possible to have $z_{i,t}$ as composed by some intrinsic value and the fixed effect: $z_{i,t} = z*_{i,t} + \mu_i$. Subtracting the variables for consecutive years can remove the fixed effect μ_i. By utilizing differential Moran scatter plots and by calculating the differential Moran's I the robustness of the conventional Moran's I can be evaluated.

3.5 The Getis Filter

From early studies in the regression framework of convergence analysis, it has been assumed that each region develops independently and reaches its long-run

[2] the variable is transformed so that the average is 0 and the standard deviation is 1

equilibrium state. However, recent studies have tried to deal with the spatial independence "hypothesis" by including spatial models when doing regressions; for example, two of the first studies to include spatial effects in growth regressions are Rey and Montouri (1999) and Bernat Jr (1996). A different approach for dealing with strong spatial autocorrelation signs is to filter the variables and express them as the sum of a spatially dependent and a spatially independent components.

One of the filters reported in the literature is known as the Getis filter Getis and Griffith (2002). This approach requires of four steps: 1) evaluating the local G spatial statistic for each region for a number of increasing distances; 2) finding the distances d_i for which the statistic decreases for each region i; 3) computing the mode d_m of the distribution of d_i; and 4) filtering the variable x_i for each region using the following equation:

$$x_i^* = \frac{x_i(W_i)}{(n-1)G_i(d_m)} \tag{5}$$

Where x_i^* is the filtered variable, n is the number of observations, W_i is the sum of all geographic links usually weighted so that is just includes the neighbors which are at closer distances than d_m , and $G_i(d_m)$ is the statistic of Getis and Ord (2010).

It is necessary to clarify that the spatial filtering can be utilized only is several restrictions on the variables are fulfilled, so that as Getis and Griffith state "The use of this approach is limited by the nature of the 1992 version of the $G_i(d_m)$ statistic. All variables must have a natural origin and be positive" Getis and Griffith (2002). In our case the variable to be studied is RGDP per capita which is positive and have a clear natural origin and thus can be decomposed using the Getis filter.

4 Results

4.1 Distribution dynamics of the original variable

A first approach to study regional mobility is to calculate the non-parametrical density estimate of the income distribution. Even though there may be merits in studying the distribution of the absolute income per capita, it is preferred to analyze the distribution of the relative income per capita. Studying the variables in relative terms allow us to remove the global (national) effect and focus on the local (district) effect. The nonparametric densities are calculated using a Gaussian kernel for the positive interval, the plots of the densities in the time t and $t+5$ (that means all 5 year transitions in the 2010-2017 sample) are shown in figure 3.

In this framework the presence of more than one mode indicates lack of global convergence while the tendency towards uni-modal distributions indicate signs of convergence. In figure 3 it can be seen that the location of the main mode has shifted to the right over the 5 year transition, indicating that the poorest regions have moved closer to the average of the entire (514 districts) distribution. It is not clear that the distribution is transitioning towards unimodality given that in time t two lumps in the distribution appear between 1 and 3 and by the time $t+5$ those two modes are still present though closer to 1 (see dashed line).

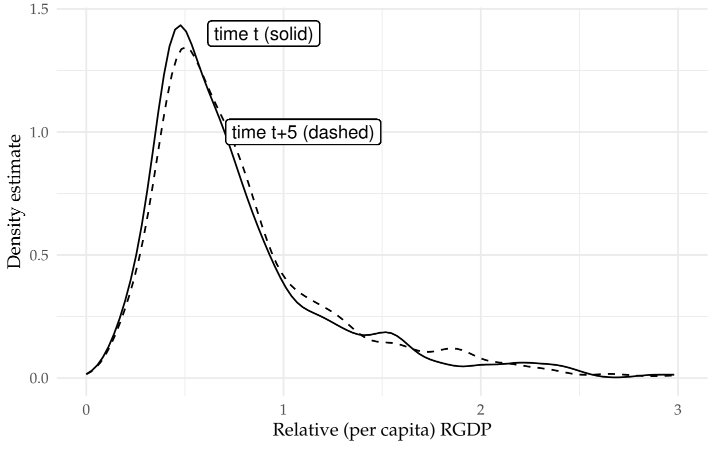


Fig. 3: Density estimate of relative per capita income at time t and t+5

Notes:(The original sample includes 514 districts. the density is shown districts for which the relative per capita income is between 0 and 3.

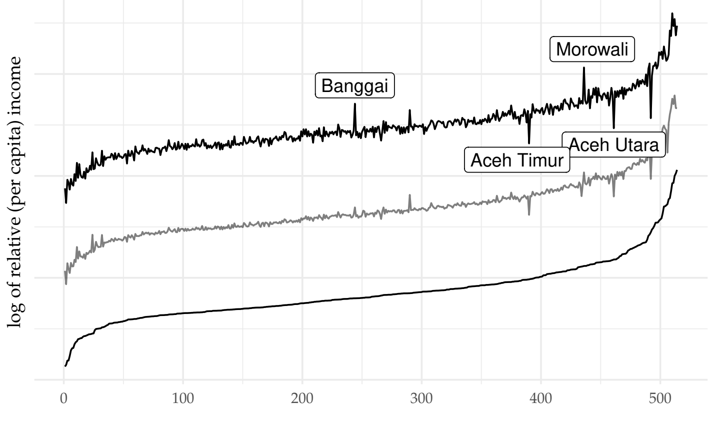


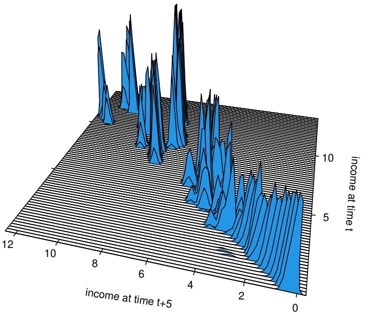
Fig. 4: Cross-profile dynamics across 514 Indonesian districts, retaining the ranking fixed at the initial year, relative (per capita) income, advancing upwards: 2010, 2013 and 2017

In order to study intra-distribution dynamics a cross profile of the regions is shown in figure 4. In the y axis the log of the relative income per capita is shown while in the x axis the regions are ranked from the lowest to the highest RGDP per capita in the year 2010. The ranking is left fixed as given for the initial year (2010), the curve for this years is the one in the lowest position. Moving upwards, the curves for 2013 and 2017 are shown. In this plot if inta-distributions mobility were null, the two upper curves would be given by monotonically increasing functions. However, that is not case and the jaggedness of the curves increases over time, showing strong signs of intra-distribution mobility.

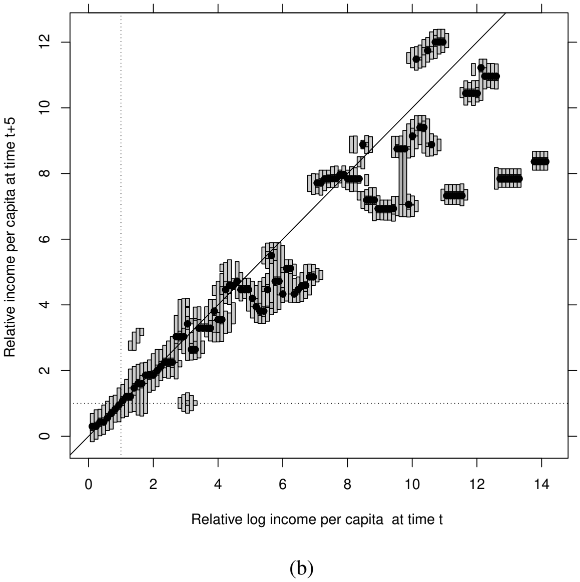
Moreover, in figure 4 the curve representing the cross-section in the year 2017 presets several local peaks; some of them include the ones of the following districts: Banggai and Morowali. These peaks represent regions which after the 7-year period ended up with higher income per capita than regions with similar initial rankings. In contrast, regions such as Aceh Utara or Aceh Timur experienced significant declines in their income per capita when compared to regions with similar income in the initial year. Figure 4 is useful in showing the local transitions of individual regions over time. For studying the overall pattern of transitions of the data we proceed to compute the conditional density.

Regional inequality at the district level in Indonesia is extremely high. For example, in the year 2010 some of the regions were 14 times richer than the average region. The conditional density is estimated for a transition period of 5 years using equation 4[3]. The results are shown in figure 5. In panel (b), The lighter shaded areas in each strip is a 99% HDR, and the darker shaded areas a 50% HDR. The mode of each conditional density is the dark point on each strip.

[3] The income data is available for the years 2010 to 2017. Considering transitions of 5 years, the mapping is the "average" of the pooled transition between 2010, 2011 and 2012 into 2015, 2016 and 2017.



(a)



(b)

Fig. 5: Relative income dynamics across 514 Indonesian districts (a) stacked density plot, and (b) highest density regions boxplot

Notes: conditional density is estimated over a 5-year transition horizon between 2010-2017. Estimates are based on a Gaussian product kernel density estimator. TThe stacked conditional density plot and the high density region boxplot were estimated at 80 and 120 points, respectively. Calculations of the plots were performed using the R package HRDCDE.

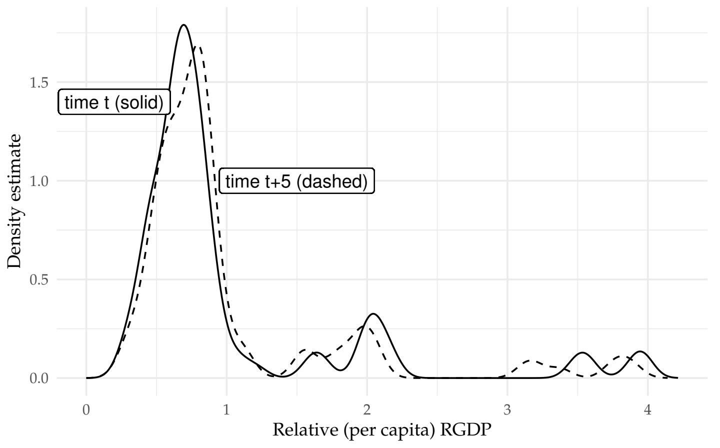


Fig. 6: Province level density estimate of relative per capita income for the initial and final times of a 15-year transition

Notes:The original sample includes 34 provinces. 5 year transitions considered for the sample of years 2010-2017

The diagonal in panel (b) represents the stagnation line, that is regions that on average did not change their relative positions over the 5-year transition period would be represented by modes along this line. Moreover, in panel (b) there are many modes located below the stagnation line for relative initial income over 4. For this districts their relative income decreased over the 5-year period. In the initial part of the distribution most modes are located on the stagnation line. In contrast, some modes between 10 and 12 are located above the stagnation line, which represents a relative increase in their income per capita.

In this framework, convergence clubs are formed when modes are located on a line parallel to the x axis in the boxplot (panel (b) in figure 5). The presence of at least 3 convergence clubs can be seen in panel (b), a lower club between 4 and 7, an upper club between 8 and 14 and the highest club for the regions with the relative log of income between 10 and 13, though relatively smaller.

A similar analysis can be done for the income per capita at the province level. One drawback of studying a lower dissagregation level is that the number of regions is dramatically reduced. There are in total 34 Indonesian provinces. In terms of l inequality, the distribution of provinces is more compact than the one for districts as mentioned in the previous paragraphs. In the year 2010 East Kalimantan is the province with the higher income per capita, but in relative terms is about 4 times richer than the average province. A density estimate of the relative income per capita for the time t and t+5 is shown in figure 6.

The density estimates in figure 6 show multimodal distributions for the initial and final times. It is clear by looking at the range of the curves that the distrubution has compressed over the 5-year transition. Moreover, the highest mode has slightly

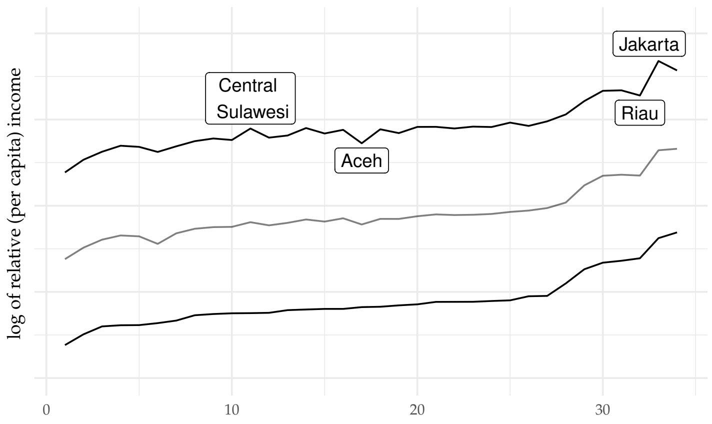
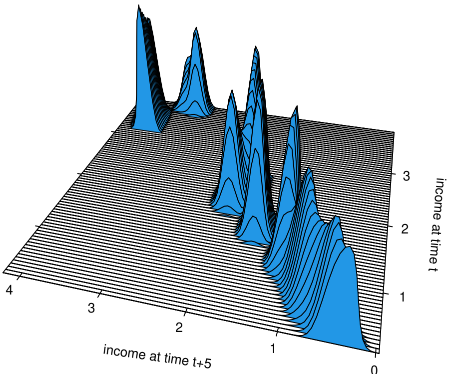


Fig. 7: Cross-profile dynamics across 34 Indonesian provinces, retaining the ranking fixed at the initial year, relative (per capita) income, advancing upwards: 2010, 2013 and 2017

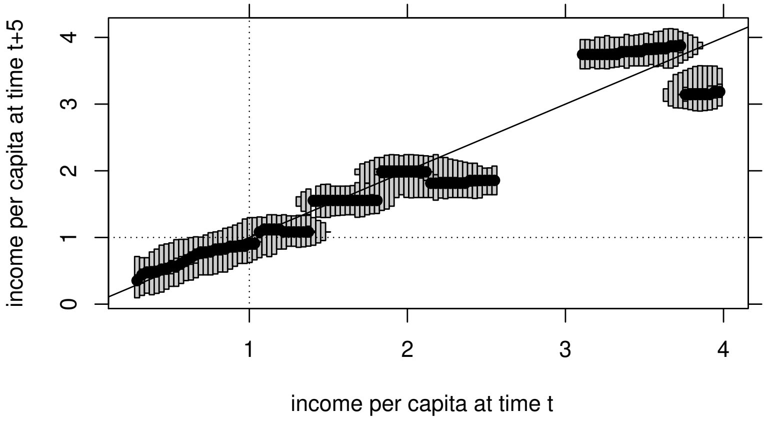
moved towards the middle of the income distribution. For GRDP per capita higher than the average (1) there are many low modes for both years, these are likely to be outliers and moreover appear to create single modes because the initial sample of regions contains a small number of regions, that is 34 provinces. Such low "gaussian-like" modes are not seen in the case of districts in figure 3. A puzzling feature in Figure 6 is that a second lower starts to appear next to the major mode at around 50% of the average income for the year $t+5$. This second mode implies that a formation of a bimodal distribution may take place in coming years. From the density plots, signs of polarization and convergence are both shown, this will also be encountered when studying the distributional convergence of the relative income per capita of the Indonesian provinces.

In figure 7 the income profile of the cross-section of provinces is shown. In this graph the ranking of the regions from the poorest to the regions has remained fixed. As discussed previously, given that the profile in the years 2010 and 2017 does not present a monotonically increasing function it is safe to assume that there has been significant intra-distrubution mobility. For example, the provinces of Central Sulawesi and Jakarta have outstandingly outperformed regions with a similar initial income per capita. In contrast Aceh and Riau have performed poorly over the 7-year time frame and their income has reduced considerably when compared to its neighbors in the initial ranking. The most extreme case is seen in Aceh; ts economy transitioned from being the 17th wealthiest in 2010 to being among the 8 poorest in 2017.

Next, in figure 8 the stacked density plot and the bloxplot for the highest density regions is shown. In panel (b) the formation of convergence clubs appears clearly.



(a)



(b)

Fig. 8: Relative income dynamics across 34 Indonesian provinces (a) stacked density plot, and (b) highest density regions boxplot

Notes: conditional density is estimated over a 5-year transition horizon between 2010-2017. Estimates are based on a Gaussian product kernel density estimator. The stacked conditional density plot and the high density region boxplot were estimated at 80 and 120 points. Calculations of the plots were performed using the R package HRDCDE.

there seem to be 3 convergence clubs as seen by the approximately 3 horizontal lines that connect the modes of the distributions. The location of the two upper clubs in the right side of panel (b) present counter-intuitive patterns. On the one hand, for a region located above 3.7 times the mean income it appears that after the 5 year transition its income is reduced to about 3 times the average (see the y axis in panel (b)). On the other hand regions that have from 3.2 to 3.8 times the average income in time t are more likely to converge to around 4 at time $t+5$. These patterns present a similar idea to what is seen as a group of local minimums and maximums at the upper end of figure 7. The lower club is located between 1.5 and 2.5.

To summarize, there appear to be 3 convergence clubs in the income distribution for provinces in Indonesia. a lower clubs near the mean and two upper clubs which present highly mobile regions. That is, the richest regions are volatile and their income ranking appear to be less stable when compared to the poorest provinces.

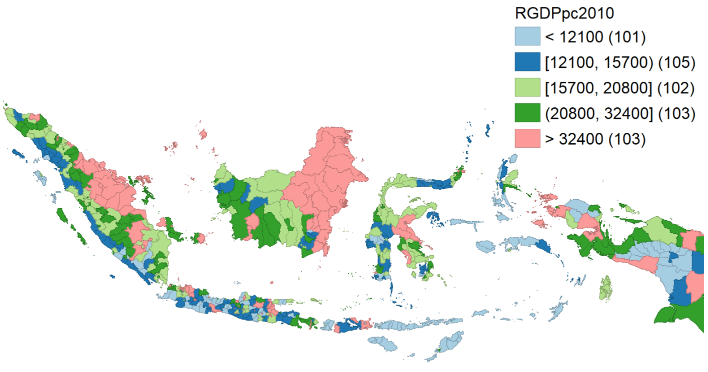
4.2 Global Spatial Autocorrelation and Spatial Patterns

From a purely visual standpoint, the income gaps across Indonesian districts can be observed in figure 9. In panel (a) the 2010 choropleth map is divided in quantiles so that each bin has approximately the same number of districts (about 103). In order to see the development progress the same breaks are plotted in panel (b) for the year 2017. Overall, there are fewer than 100 regions in the first to bins in the year 2017, this means that several districts escaped lower income groups in 2010, joining higher income clusters in 2017.

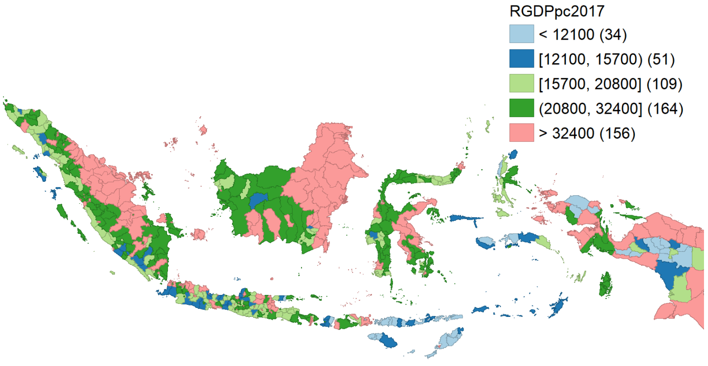
Moreover, the Regions in east Borneo remained over time in the top income groups. Similarly regions in the north-eastern part of Sumatra are also in the highest income category in 2010 and 2017. In addition, the other districts surrounding the pink clusters on those island have also developed and join the second highest income group (in dark green). A comparable process takes place in central Java where regions that where in the lowest and second lowest income groups (light and blue groups) have increased their per capita income levels and joined the 2nd and 3rd highest income groups.

Eastern Indonesia regions are mostly characterized by low income districts, most low income regions in 2017 can be found there. A shocking fact is that throughout the 7-year period many regions int he periphery of West Papua have dramatically increased their economic output, yet there is a core of regions that remain in the lowest income in both 2010 and 2017. To sum up, from both choropleth maps it can be concluded that regions with similar output per capita are spatially clustered. We proceed to present a formal evaluation of spatial dependence.

As explained in the methods section, the application of the Getis filter requires that the data shows strong signs of spatial autocorrelation. In the case of provincial data (34 provinces) it was found that there are no signs of spatial autocorrelation for Regional GDP (RGDP) per capita for any year in the period 2010-17, with the Moran's I statistic being not being statistically different from 0. These results do not change whether we use a distance band or a queen contiguity criteria when creating the spatial weights matrix.



(a)



(b)

Fig. 9: Income per capita at the district level in Indonesia in 2010 and 2017

Notes: The quantiles were created for the year 2010 and thh same breaks are plotted for 2017.

In contrast, for district level data, using a queen contiguity criteria for the Thiessen polygons, it turns out that the Moran's I is significant with a $p - value < 0.01$ for all years from 2010 up to 2017. However, by using the differential Moran's I it is found that this latter statistic is robust from 2013. As it is common in the literature the value of the differential Moran's I is significantly smaller that the one reported for the standard Moran's I.

In Figure 10 the magnitude of both Moran's I statistics is shown. In addition, it can be stated that the standard statistic increased from 2010 to 2017 until reaching a

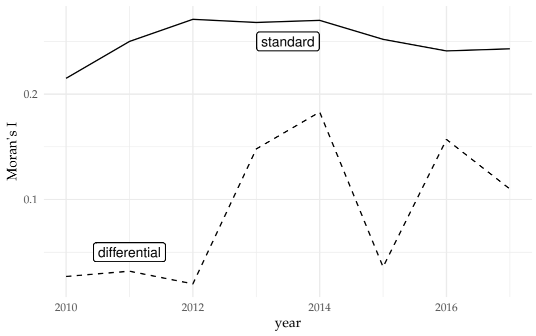


Fig. 10: Differential and standard Moran's I at the district level

point at around 0.25. Moreover, in the case of the differential statistics the pattern is more erratic being significant from 2013.

4.3 The Spatially filtered variable and its dynamics

As shown in the previous section strong signs of spatial autocorrelation are present for districts in Indonesia (see figure 10). Therefore, in order to use the Getis filter we proceed to find the critical distance d_m. Following the process explained in the methods section, the critical distance found was $d_m = 338km$. Using this distance and plugging it in the equation 5; the spatially filtered component of Regional GDP per capita for each location and year can be found.

The density estimate of the spatially filtered per capita income for the years 2010 and 2017 (in 5 year transitions) is shown in figure 11. Similar to figure 4, the major mode of the distribution in figure 11 shifts to the right over the 5-year period. In addition, there seem to be at least 3 modes in both years; located at around 0.6, 1.2 and 2 in times t and $t+5$, respectively. There seem to be a significant group of regions for which the spatially filtered income per capita is a very small fraction of the average, as the main mode is well below 1 in both initial and final years. This will be also encountered when estimating the conditional density.

One strategy to evaluate the presence of spatial effects is to analyse the data by using the distributional dynamics framework. Although, this analysis is usually done for different years as explained in the methods section of this paper, instead of using the transformation for different years we study the mapping form the original variable to the filtered non-spatial variable as set by equation 5. The modal box plots for the initial year t shown in Figure 12. From panel (b) in figure 12 it can be stated that the distribution is different from the trivial 45 degrees line.

It appears that the modes are located below the diagonal line for the regions in the lower part of the distribution; while at the upper end modes are located above the persistence line. The complexity of the transformation as shown in figure 12, does not clearly suggest whether the filtered distribution is more disperse, though it appears

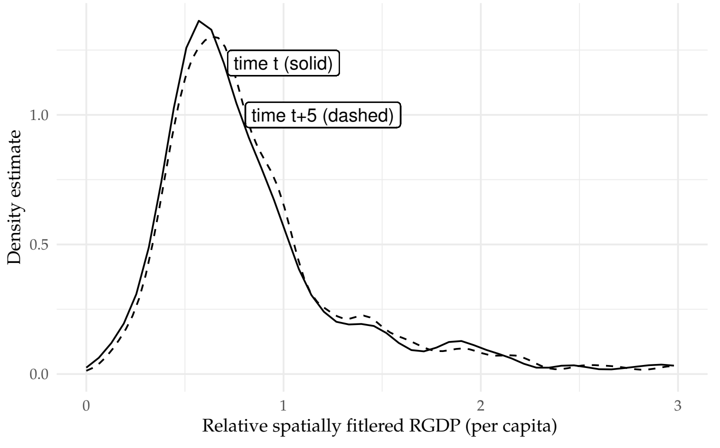


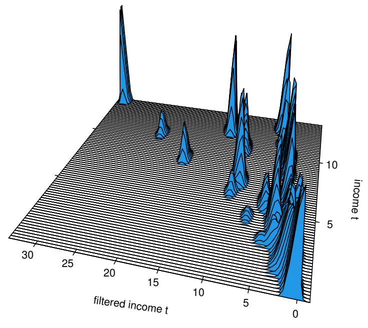
Fig. 11: Density estimate of the spatially filtered per capita income for 5-year transitions in the period 2010-2017

Notes: The original sample includes 514 districts. the density is shown for districts for which the relative filtered per capita income is between 0 and 3.

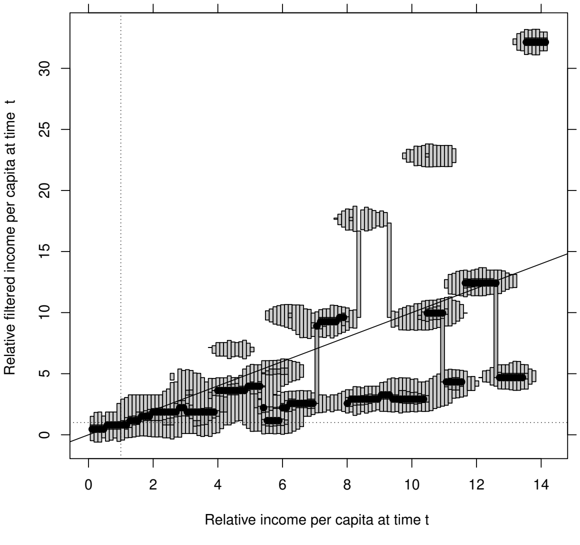
that the filtered data has larger gaps, as some regions have more than 30 times the average filtered income. This will be further analyzed in the following paragraphs.

A dynamical analysis of the filtered variables is also presented. Considering a 5-year transition period, the stacked density plot and the modal bosxplots are presented in figure 13. The filtered data shows that the least developed regions are located mostly over the 45 degrees line. Moreover, the regions in the middle of the distribution are mostly located over or near the diagonal line, showing persistence patterns for these regions. In contrast, there appears to be a convergence club in the upper tail of the distribution; districts with relative income between 15 and 30 in the initial year are likely to end up with a value close to 16 5 years later. Overall, it can be seen that the filtered data shows strong convergence patterns in the upper tail of the distribution.

Overall, the spatial effects seem to be far from trivial. For the initial time in figure 12 the filtered data shows both polarizing and blending patters; relatively similar regions in terms of original income being mapped to both high and low filtered variables, that is over and below the stagnation line. Furthermore, comparing the transitional period of 5 years for the original and filtered variables, that is figures 5 and 13, the differences seem to appear at the upper end of the distribution. For the lowest income regions the spatially filtered data shows that after the 5 year period the modes are located above the persistence line. The reduction in inequality can be seen in figure 13 for the upper quartile of the initial distribution (x axis).



(a)



(b)

Fig. 12: Stochastic kernel mapping from the original to the spatially filtered distribution across 514 Indonesian districts (a) stacked density plot, and (b) highest density regions boxplot

Notes: conditional density is estimated over a 5-year transition horizon between 2010-2017. Estimates are based on a Gaussian product kernel density estimator. The stacked conditional density plot and the high density region boxplot were estimated at 80 and 120 points, respectively. Calculations of the plots were performed using the R package HRDCDE

We proceed to evaluate the dynamics of the filtered and original variables trough the lenses of the classical convergence framework. This analysis will shed more light on the polarizing or otherwise merging processes that may be related to spatial effects.

4.4 Classical convergence framework

We now proceed to evaluate the convergence using the classical framework, for this we compare the growth rates of the filtered and original variables and their initial state at different administrative levels using the equation 1. The beta convergence plots for districts are shown in figure 14. The slope in panel (b) is steeper than in panel (a), which means that the speed of convergence is larger for the filtered variable. On average for both variables, the poorest districts seem to be growing at a faster pace that the richest regions. In fact, these differences can be appreciated in table 1. In this table the half-life is also included; given the higher speed, the half-life is reduce by 20% when considering filtered variables.

Finding signs of Beta convergence (and sigma convergence) at the district level is in itself a telling and remarkable result about the development of Indonesia. The evidence for the convergence of such dis-aggregated data in many other developing nations is far from trivial, and convergence is found in just a small number of samples. Speeds of convergence of about 1% (0.9% for income and 1.1% for the spatially filtered variables) at the district level is lower than the 1.4% reported at the provincial level. The provincial level speed is not very different from the one found by Barro and Sala-i Martin (1992b) for prefectures in Japan and states in the USA of about 2%. Which is very common in the regional convergence literature[4].

Table 1: Classical convergence framework summary

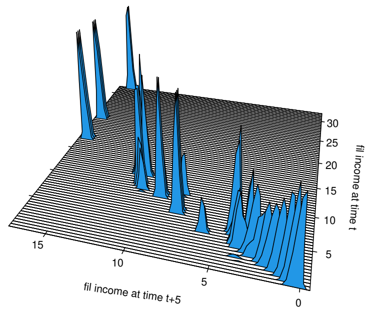
data (spatial scale)	**speed of convergence**	**half-life(years)**	$quotient(\sigma)$
RGDPpc(districts)	0.93***	74.2	1.03
RGDPpc(district-filtered)	1.14***	60.6	1.04*
RGDPpc(provinces)	1.44***	48.6	1.06

Note: unity. The speed of convergence is derived from a linear regression between the growth rate and the initial income value. *, **, *** indicate significance at the 10%, 5%, 1% levels respectively of the regression coefficient. For the quotients, the significance refers the difference from the null hypothesis of a quotient equal to 1.

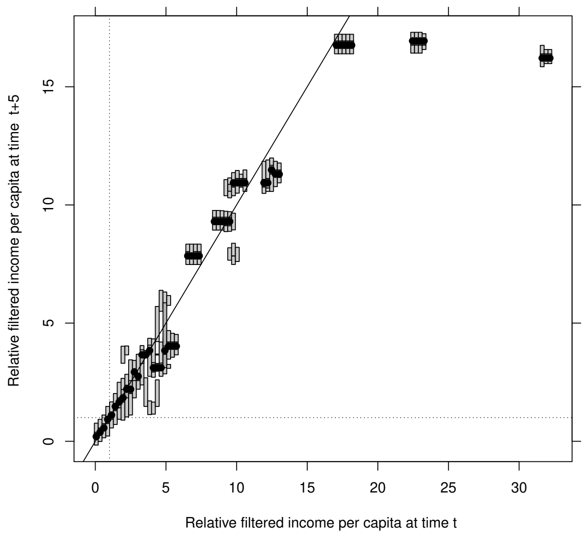
Moreover, after concluding that beta convergence signs are robust for both variables, it remains to be seen what occurs in terms of sigma convergence. The quotients (Initial value / final value) in table 1 at both spatial levels are larger than 1. Thus, weak signs of sigma convergence are present for both the original variable and the non-spatial variable using both statistical measures. However, sigma convergence is significant for the filtered variable only as seen by the significance of 10%.

Thus, Spatial effects seem to be responsible for a slower beta and sigma convergence, given that both are faster for the filtered variables. In addition, it seems

[4] The beta convergence plot for provinces can be found in the appendix



(a)

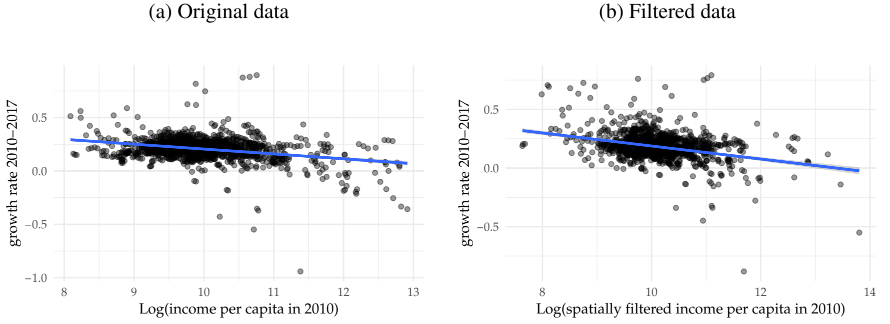


(b)

Fig. 13: Stochastic kernel mapping for the spatially filtered distribution across 514 Indonesian districts (a) stacked density plot, and (b) highest density regions boxplot

Notes: Estimates are based on a Gaussian product kernel density estimator. The stacked conditional density plot and the high density region boxplot were estimated at 80 and 120 points, respectively. Calculations of the plots were performed using the R package HRDCDE.

Fig. 14: Convergence of RGDPpc at the district level



Notes: The solid line indicates the fit of a linear regression.

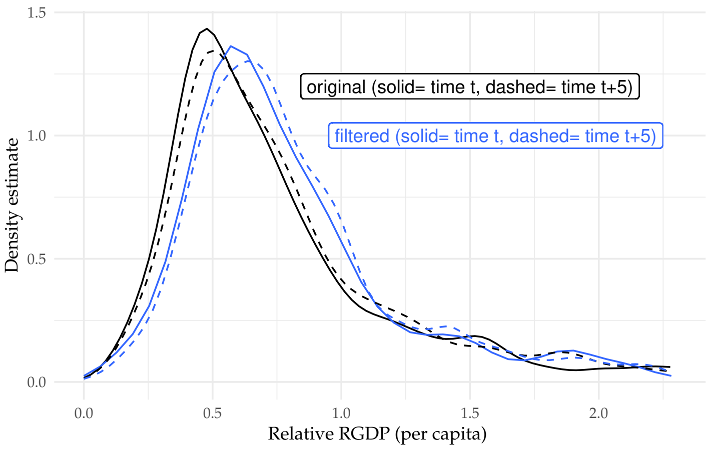


Fig. 15: Density estimate of relative per capita income at time t and t+5, for original and filtered data

Notes:(The original sample includes 514 districts. the density estimate was truncated at x=2.2

that such effects are increasing over time, likely related to the fact that spatial autocorrelation has also increased in magnitude over time.

On the one hand, spatial effects seem to be responsible for the polarization of data in the sense of the formation of upper and lower convergence clubs in figure 12. On the other hand, neighboring effects also seem to be correlated with the overall slower speed of convergence of 0.93% instead of the non-spatial 1.14%.

Similar results for the spatial effects can be observerd in figure 15. In blue in this figure the density estimates of the filtered data are plotted and in black the estimates for the original income data. The modes of the blue curves at time t and more so at time t+5 are near 1. Therefore, the filtered income data displays smaller disparities as more regions have income values closer to the mean.

Overall, it appears that spatial effects are slowing the convergence process in Indonesia. Moreover, the filtered data appear to have larger values both in terms of the mean and the median of income. Therefore, we conclude that spatial effects are playing an overall negative effect in the regional development of Indonesia.

5 Concluding Remarks

The issue of whether poor districts are converging to the level of richer districts after the implementation of decentralization is an important concern for policymakers in Indonesia. This study has pointed out two results for policy discussion.

From a policy standpoint, the results imply that the policies and capabilities of local governments to support income distribution programs remain ineffective. Therefore, the role of the central government in reducing inequality, mainly through affirmative policies to poor regions, is still needed. Moreover given the tremendous role that space is playing (and even an increasing role given the increments in the Moran's I measures) in the development of Indonesia, it seems safe to say that breaking spatial barriers and the infrastructure and capabilities required for it must be on top of the policy agenda.

The transition dynamics of the unfiltered distribution and the filtered distribution indicates a high degree of mobility at the top of the income distribution (convergent mobility is accomplished in terms of lower relative income for the richest regions). However, by also incorporating a classical convergence analysis of the spatial and non-spatial variables, we were able to conclude that spatial effects are slowing the convergence process. Thus, the findings contrast the previous analysis such Vidyattama (2013) that show a significant role of neighbor effects in reducing inequality in Indonesia during the decentralization period.

The contribution of this study is three fold. First, the construction of a balanced panel dataset and the geo-referenced location of the capital cities of the districts allows us to analyse the patterns of all regions in Indonesia in post-decentralization. Second, following Fischer and Stumpner (2010), in order to avoid misleading interpretations on evaluating cross-regional growth and convergence dynamics, the analysis were also conducted using spatially filtered data. The filtered data calculated by the Getis spatial filter help us evaluate the convergence dynamics across districts more precisely and find the patterns of growth for the poorest regions, which are per se of high-priority for policy implementation. Lastly, it was argued that the neighboring effect is in fact reducing the speed of convergence for the cross-sectional distribution of districts.

The construction of a balanced dataset for income per capita of provinces and districts from 2010 up to 2017 can be very useful for further research. For example, in terms of convergence studies, a time series convergence approach as suggested by Phillips and Sul (2009, 2007) can be performed in order to evaluate the temporal dynamics of the cross-section of provinces and districts. Also, the shapefile together with income data can be used in order to study local spatial patterns, by analysing the local local Moran's I and other measures of local spatial autocorrelation (Anselin (1995)). This type of study can help us to find clusters of regions with similar growth performance and uncover more details about the neighboring effects documented in this paper.

References

Aginta H, Gunawan AB, Mendez C (2020) Regional income disparities and convergence clubs in indonesia: New district-level evidence 2000-2017. MPRA Working Paper 99079

Akita T (1988) Regional development and income disparities in indonesia. Asian Economic Journal 2(2):165–191

Anselin L (1995) Local indicators of spatial association-lisa. Geographical analysis 27(2):93–115

Anselin L, Syabri I, Kho Y (2006) Geoda: an introduction to spatial data analysis. Geographical analysis 38(1):5–22

Aritenang AF (2014) The spatial effect of fiscal decentralisation on regional disparities: the case from indonesia. Indonesian Journal of Geography 4(1):1–11

Barro R, Sala-i Martin X (1991) Convergence across states and regions. Brookings papers on economic activity 1991(1):107–182

Barro R, Sala-i Martin X (1992a) Convergence. Journal of Political Economy 100(2):223–251

Barro R, Sala-i Martin X (1992b) Regional growth and migration: A japan-united states comparison. Journal of the Japanese and International Economies 6(4):312–346

Barro RJ (1991) Economic growth in a cross section of countries. The quarterly journal of economics 106(2):407–443

Barro RJ, Sala-i Martin X (1995) Economic growth / Robert J. Barro, Xavier Sala-i-Martin. McGraw-Hill New York

Bernat Jr GA (1996) Does manufacturing matter? a spatial econometric view of kaldor's laws. Journal of regional science 36(3):463–477

Cravo TA, Resende GM (2013) Economic growth in brazil: a spatial filtering approach. The annals of regional science 50(2):555–575

Esmara H (1975) Regional income disparities. Bulletin of Indonesian Economic Studies 11(1):41–57

Fingleton B (2004) Regional economic growth and convergence: insight from a spatial economic perspective. In: LAnselin RF, Rey S (eds) Advance in Spatial Econometrics, Berlin: Springer-Verlag

Fischer MM, Stumpner P (2010) Income distribution dynamics and cross-region convergence in europe. In: Handbook of Applied Spatial Analysis, Springer, pp 599–628

Garcia JG, Soelistianingsih L (1998) Why do differences in provincial incomes persist in indonesia? Bulletin of Indonesian Economic Studies 34(1):95–120

Getis A, Griffith DA (2002) Comparative spatial filtering in regression analysis. Geographical analysis 34(2):130–140

Getis A, Ord JK (2010) The analysis of spatial association by use of distance statistics. In: Perspectives on Spatial Data Analysis, Springer, pp 127–145

Hendajany N, Saepudin D, Suaesih A (2017) Beta convergence and sigma convergence : Evidence from indonesian district-level data. Jurnal Ekonomi Kuantitatif Terapan 10(2):19–25

Hyndman R, Bashtannyk D, Grunwald G (1996) Estimating and visualizing conditional densities. Journal of Computational and Graphical Statistics 5(4):315–336

Hyndman RJ (2019) hdrcde: Highest Density Regions and Conditional Density Estimation. URL http://pkg.robjhyndman.com/hdrcde, r package version 3.4

Kataoka M (2012) Economic growth and interregional resource allocation in indonesia. Studies in Regional Science 42(4):911–920

Kurniawan H, de Groot HL, Mulder P (2019) Are poor provinces catching-up the rich provinces in indonesia? Regional Science Policy & Practice 11(1):89–108

Magrini S (2009) Why should we analyse convergence using the distribution dynamics approach? Scienze Regionali 8(1):5–34

Mendez C (2017) Heterogeneous growth and regional (di) convergence in bolivia: A distribution dynamics approach. Coyuntural Economics 2(4):81–108

Mendez C (2018) On the distribution dynamics of human development: Evidence from the metropolitan regions of bolivia. Economics Bulletin 38(4):2467–2475

Moran PA (1948) The interpretation of statistical maps. Journal of the Royal Statistical Society Series B (Methodological) 10(2):243–251

Pepinsky TB, Wihardja MM (2011) Decentralization and Economic Performance in Indonesia. Journal of East Asian Studies 2011(11):337–371

Phillips PC, Sul D (2007) Transition modeling and econometric convergence tests. Econometrica 75(6):1771–1855

Phillips PC, Sul D (2009) Economic transition and growth. Journal of Applied Econometrics 24(7):1153–1185

Quah D (1993a) Empirical cross-section dynamics and tests of the convergence hypothesis. European Economic Review 37:426–434

Quah D (1993b) Galton's fallacy and tests of the convergence hypothesis. The Scandinavian Journal of Economics 95(4):427–443, URL http://www.jstor.org/stable/3440905

Quah D (1996) Twin peaks: Growth and convergence in models of distribution dynamics. Economic journal 106(437):1045–1055

Quah D (1997) Empirics for growth and distribution: Stratification, polarization, and convergence clubs. Journal of Economic Growth 2(1):27–59

Rey SJ, Gallo J (2009) Spatial analysis of economic convergence. In: Mills TC, Patterson K (eds) Palgrave Handbook of Econometrics Volume 2: Applied Econometrics, London: Palgrave Macmillan UK, pp 1251–1290

Rey SJ, Montouri BD (1999) Us regional income convergence: a spatial econometric perspective. Regional studies 33(2):143–156

Talitha T, Firman T, Hudalah D (2019) Welcoming two decades of decentralization in Indonesia: a regional development perspective. Territory, Politics, Governance 17(17):1–19

Vidyattama Y (2013) Regional convergence and the role of the neighbourhood effect in decentralised indonesia. Bulletin of Indonesia Economic Studies 49(2):193–211

Villaverde J, Maza A (2012) Chinese per capita income d istribution, 1992–2007: a regional perspective. Asian Economic Journal 26(4):313–331

Wieland T (2019) Reat: A regional economic analysis toolbox for r. REGION 6(3):R1–R57

6 Appendix: Districts Interpolation

6.1 Districts Interpolation

The Indonesia Central Bureau of Statistics has published periodically the data of per capita regional gross domestic product (Per Capita Regional GDP) at district (regencies and cities) level at current price. However, the sequential time series database of Per Capita Regional GDP at constant price at districts level are difficult to obtain. Even if the data are available, the number of districts has been increasing over time due to the proliferation of administrative provinces and districts from the beginning of the decentralization period. The number of districts increased significantly from 341 districts in 2000 to 514 districts in 2017 (See Figure 16). Besides that, the year basis for calculating the constant price has changed several times. Therefore, in this paper, we construct the full balanced panel data for 514 districts level in Indonesia.

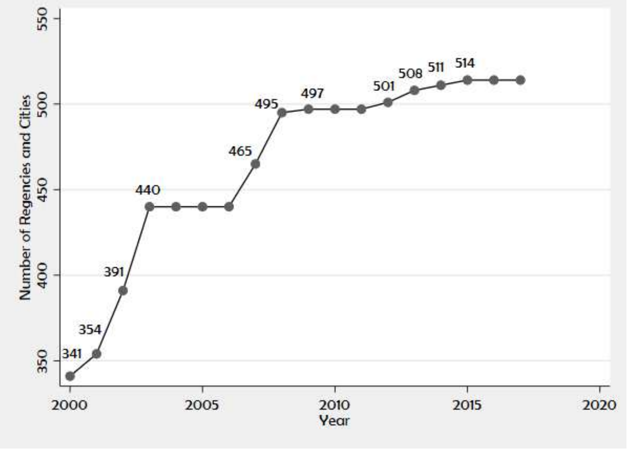


Fig. 16: Number of Districts in Indonesia since 2000

In this paper, we construct the Regional GDP per capita at constant price by calculating Gross Domestic Product (GDP) at constant price and population at all districts. To do so, we conducting interpolation or imputation for all missing value in both the Gross Domestic Product (GDP) at constant price and the number of populations. The interpolations were conducted to all missing values in 173 districts during the period of 2000 to 2017. In addition, we also made adjustment (replacement) on the historical data in some reference districts (origin districts), mainly in the period before they split-up into the other new district(s). This adjustment is conducted in order to avoid the structural break of the time series, i.e. the time series of reference districts before and after split-up.

To conduct the imputation, this paper uses a linear regression method with the reference districts and year as regressors, similar methods conducted by Kurniawan et al. (2019). The origin district is chosen as reference for imputation. We assume that the new district is having co-movement with its reference district. Since the all districts have their actual data at least since 2015, then the trend of the new districts refers to their actual data in the years when the data available and follow its reference district in the year when they are interpolated.

To illustrate the imputation method in this study, we use one of the proliferated districts and its reference district in Province of Jambi, which are the city of Sungai Penuh City and the regency of Kerinci. The city of Sungai Penuh separated from its origin district, the Regencey of Kerinci, and became a new district in Oktober 2009. Therefore, data for the City of Sungai Penuh before 2010 are not available and the data for the Regency of Kerinci after 2010 will be drop significantly due to the administrative creation of new district. In order to solve missing data and the structural break due to the administrative creation of new district, this study conduct the imputation/interpolation as follows.

6.2 Constructing Reference District and Interpolated the New District Data

The historical data of origin district before split-up and the sum of the composite of new district(s) data after the proliferation are used as reference district. In Figure A1 (in figure 17) , the reference district is the combination of the GDP data of Kerinci Regency in 2000-2009 before proliferation and the sum of the GDP data of Regency of Kerinci and City of Sungai Penuh after the proliferation in 2010-2017. This paper uses this reference data to conduct the imputation for the missing data of Sungai Penuh City before proliferation. The result can be seen in Figure.A1 (in figure 17) .

6.3 Data Adjustment for Origin District

Since the origin district were split-up into new district, the time series data may have structural break, especially during the proliferation year. Therefore, the time series data of origin district are adjusted by subtracting the time series data of reference district with the imputed data of new district before the proliferation year. Basically, this calculating will give the similar pattern (slope) of the data, but in different intercept to the origin district, see Figure A2 (in figure 17). The comparison of imputed GDP data for Kerinci Regency and Sungai Penuh City can be seen in Figure.A3 (in figure 17) and the sum of both districts is equal with the reference districts.

6.4 Constructing Population Data

The missing data of total population in each district are imputed using linear calculation of the share of population to total province population. This paper uses the population data in each district from INDODAPOER-World Bank and population data at province level from Central Bureau of Statistics as a basis data for interpolation.

6.5 Calculating Per Capita GDP Data

Since the interpolation for both Regional GDP and total population at district level are done, the next step is calculating Per capita GDP at constant price for all districts. Per capita GDP at constant price for districts level are obtained by dividing Regional GDP (as numerator) by the total population (as de-numerator) in each district. Both the database for Regional GDP and Per Capita Regional GDP at districts level in this paper have been adjusted using the latest constant price 2010. As an example, the comparison of Per Capita Regional GDP at districts level between Kerinci Regency and Sungai Penuh City can be seen in Figure A.4. (in figure 17).

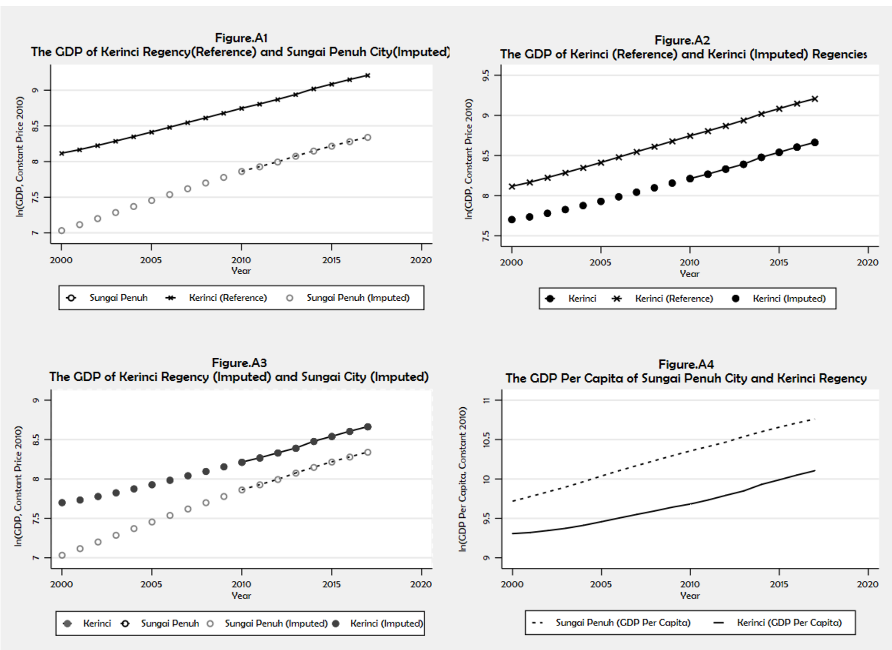
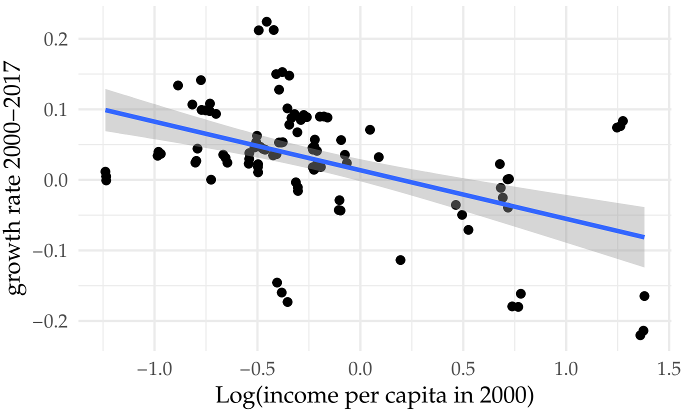


Fig. 17: Comparison of Imputed Data and Its Reference

This study realizes the measurement error that can be exist caused by the imputation/interpolation approach. However, this study uses actual data of the proliferated districts and reference districts, particularly the data after 2011 that have already used the constant price 2010 as price basis. In addition, this study also carefully compares the imputed data of proliferated districts with its reference districts.

6.6 Beta convergence at the provincial level

Fig. 18: Convergence of RGDPpc at the provincial level



Notes: The solid line and its associated confidence interval indicate the fit of a linear regression.