Regional Income Disparities and Convergence Clubs in Indonesia: New District-Level Evidence 2000-2017

Harry Aginta (✉ harry_ag@bi.go.id)
Nagoya Daigaku
https://orcid.org/0000-0002-0565-6207

Anang Budi Gunawan
Badan Perencanaan Pembangunan Nasional

Carlos Mendez
GSID, Nagoya University

Keywords: regional income inequality, convergence, districts, Indonesia

Posted Date: May 21st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-29036/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Regional Income Disparities and Convergence Clubs in Indonesia: New District-Level Evidence 2000-2017

Harry Aginta *, Anang Budi Gunawan †, Carlos Mendez ‡

Abstract

Reducing regional income disparities is a central challenge for promoting sustainable development in Indonesia. In particular, the prospect for these disparities to be reduced in the post-decentralization period has become a major concern for policymakers. Motivated by this background, this paper aims to re-examine the regional convergence hypothesis at the district level in Indonesia over the 2000-2017 period. By using non-linear dynamic factor model, this study analyzes novel data set to investigate the formation of multiple convergence clubs. The results indicate that Indonesian districts form five convergence clubs, implying that the growth of income per capita in Indonesian 514 districts can be clustered into five common trends. From the lens of spatial distribution, two common occasions can be observed. First, districts belonging to the same province tend be in the same club and second, the highest club is dominated by districts with specific characteristic (i.e., big cities or natural resources rich regions). From a policy standpoint, this findings of multiple convergence clubs at significantly different levels of income allows regional policy makers to identify districts facing similar challenges. It may have meaningful implications for regional development policies, including the call of inter-provincial development policy.

JEL Codes: O40, O47, R10, R11

Keywords: regional income inequality, convergence, districts, Indonesia

*Bank Indonesia. Corresponding author: harry_ag@bi.go.id
†Ministry of National Development Planning of Indonesia.
‡Graduate School of International Development (GSID) Nagoya University, Japan.
1 Background

The goal of development progress in a country is to achieve social welfare. Equitable development does not only prevent social conflict, but is also needed as a prerequisite for achieving social justice. One indicator of the of equitable development is equality in per capita income level, both between the class community and regions. The classical economic theory developed by Solow (1956) also emphasizes a proposition concluding that developing countries will slowly converge to their developed counterparts in terms of income per capita. The premise requires a conditionality where countries with high per capita income levels grow in a lower rate in comparison with those in low per capita income country group.

Since the 20th century, however, many criticisms of the Solow theory began to emerge. Thirlwall (2001) for example, argued that it is very difficult to explain the growth difference by using the initial difference in the factor endowments. Others also argue that the underlying assumptions of the model makes it to be lacking on policy implications. Since growth is purely a function of inputs and technological progress is exogenous, domestic growth policy has no place to promote economic growth (Aghion and Howitt, 2006).

Despite many arguments concerning the convergence of a whole economy, some scholars in the field of economic growth have proved that a convergence phenomenon may exist in some economies with similar structural characteristic (Azariadis and Drazen, 1990; Galor, 1996). In other words, some economies with relatively similar structures may form a relatively common balance growth path - widely known as a club convergence hypothesis, which has been one of the growing research areas in empirical economic growth literature.

But the empirical procedure to test club convergence in fact does not seem as easy as the idea. Researchers have put lot of efforts into designing the accurate econometric methods to test the hypothesis of club convergence. Durlauf and Johnson (1995) employed regression tree analysis and found that convergence speed in income and human capital levels within a group of countries is significantly higher compared to the speed of convergence in total samples. Krugman (1999) has also examined similar issues to understand the role of geography in development and documented it in his economic geography theory.

Our study contributes on the convergence literature in following two important features. First, we conduct a novel regression by employing the log t test proposed by Phillips and Sul (2007) to identify multiple convergence clusters or clubs. In addition, we conduct the classical convergence test for each club not only to check the robustness of club convergence formation but more specifically to compare the convergence speed of each club. This is useful in particular if the goal of policy makers is to reduce income gap between regions given time constraints. Second, our analysis is based on complete district level data of per capita income in 514 Indonesian districts in the period from 2000 to 2017. We chose year 2000 as a starting period because it corresponds to the commencing of decentralization era in which state budget funds are more distributed into regions, not only to provincial government but also to municipal government, where the budget is spent into more down-to-earth programs. Hence, knowing the per capita income cluster formation among Indonesian districts in the post decentralization era is of particular relevance with regard to subsequent policies in reducing regional income inequality. Also, using more granular data like district level would enable us to uncover hidden patterns
that lying in a particular province. This is very helpful when one considers some possible inter-provincial policy outcomes.

Our results reveal that Indonesian districts form five convergence clubs, implying that the growth of income per capita in 514 districts can be clustered into five common trends. We also found the existence of "catch-up effect" within each club where lower income districts converge to higher income level. From the lens of spatial distribution, two common occasions can be observed. First, districts belonging to the same province tend to be in the same club and second, the highest club is dominated by districts with specific characteristic i.e. big cities or natural resources rich regions. Furthermore, classical convergence test provides solid evidence of convergence in each club where $\beta$ coefficient for each club is negative and significant and the fastest convergence process happens to be in Club 3 with half-life around thirteen years.

The rest of the paper is organized as follows. In Section 2, we discuss literature review on economic convergence studies, while in Section 3, the method and data used to identify convergence clubs are explained. We present our results on the formation of convergence clubs in Section 4 followed with discussion in Section 5. Finally, Section 6 concludes the paper with remarks.

2 Literature review

Recently, many studies using existing long-term data sets of Gross Domestic Products (GDP) compiled for most countries has been conducted to identify convergence clubs. The convergence itself has been seen in two dimensions, namely $\sigma$-convergence) and $\beta$-convergence (see Bernard and Durlauf (1995); Hobijn and Franses (2000); Phillips and Sul (2007)). While $\sigma$-convergence refers to change of growth dispersion (in most of cases the growth of income per capita) across countries or regions over time, $\beta$-convergence is seen in negative correlation between initial level of income capita and its growth. Implicitly, this means that low income countries tend to grow relatively faster than high income countries and thus are able to catch up (Barro, 1991; Barro and Sala i Martin, 1992). The concept of $\beta$-convergence can be differentiated into absolute and conditional convergence (see Islam (1995, 2003); Mankiw et al. (1992); Sala i Martin (1995))). On one side, in the context of absolute $\beta$-convergence, countries will approach a particular common steady-state growth path over time given the variability on each initial condition. On the other side, the notion of conditional $\beta$-convergence implies convergence occurs towards different paths of steady-state growth given some assumptions that countries have distinctive characteristics (accumulation in human and physical capital, institution, economic and political system, and other factors affecting economic growth). Indeed, many researchers found that the dispersion of income per capita across economies follow the patterns of clusters rather than the direction of commons growth path (Quah, 1996; Phillips and Sul, 2009; Basile, 2009). Not only true for largely diversified cases such as cross country analysis, but this trend has also been found in more integrated economies like those in Western Europe (Corrado et al., 2005).

In 2007, Phillips and Sul developed log t convergence test, an innovative method to investigate the existence of multiple convergence clubs based on a clustering algorithm. We choose this method because its superiority in the sense that it allows the time series not to be co-integrated, thus allowing individual observation to be transitionally divergent (Bartkowska and Riedl, 2005). The method also concludes that the absence of co-integration in respective time series does not necessarily deny the existence of conver-
gence (Phillips and Sul, 2007). Due to its advantages, this approach has been utilized by numerous researchers with the applications in convergence analysis on various economic indicators such as per capita income, financial development and energy. Against this background, this paper aims to identify convergence clubs in Indonesia to show the spatial economic outcome of the development after the year of 2000 where decentralization policy came into play. Furthermore, we will visualize the distribution of clubs on the map and provide clear explanation on the distribution formation accordingly, which is particularly paramount to understand the distribution of development geographically (Sakamoto, 2007).

Several studies on economic convergence in Indonesia indeed have been well documented (Table 1). Before decentralization era, Garcia and Soelistianingsih (1997) employed β-convergence method to investigate the existence of convergence in income per capita across Indonesian province level and found that regional income disparities tend converged and it may take between thirty to forty years to reduce income differences by half. Later, using Indonesian panel data at province level, Resosudarmo and Vidyattama (2006) found the existence of conditional convergence on provincial income per capita in the face of regional income disparities. After a decade of decentralization era, using Williamson Indexes for regional per capita income and Human Development Index (HDI), Vidyattama (2013) showed insignificant β-convergence estimates at both province and districts levels during 1999-2008 but significant β-convergence occurred in specific period of 2005-2008 at district level, partially related to the neighborhood effect. Similarly, using traditional estimation on σ and β-convergence, Tirtosuharto (2013) examined convergence in the time of crisis and aftermath. Surprisingly, they did not found convergence both during the 1997 financial crisis and the following decade from 2003 to 2012, the period when Indonesian economy slowly recovered from the crisis and the beginning of decentralization era. Recently, Kurniawan et al. (2019) employed club convergence test of Phillips and Sul (2007) and compared the convergence of per capita Gross Regional Product (GRP) with other socio-economic variables namely Gini coefficient, school enrolment and fertility rate. They found that none of those variables shows overall convergence. Instead, the development of those indicators assemble two convergence clusters. Finally, using both classical and distributional convergence frameworks, Mendez (2020) investigated convergence on regional efficiency across Indonesian provinces from 1990 to 2010. From the lens of classical convergence framework, he found the evidence of convergence in three efficiency measures, namely pure efficiency, scale efficiency and overall efficiency. The distributional convergence analysis, on the other hand, observed one local convergence cluster in scale efficiency, while overall and pure efficiency assemble two local convergence clusters. In addition, pure efficiency seems to be the driver of regional efficiency dynamics in Indonesia.

3 Methodology and Data

As mentioned before, in this paper we first identify the convergence clubs in income per capita among Indonesian districts from the year of 2000 to 2017. We extend previous study to identify convergence clubs at the district level and further test the club merging. Second, we visualize the club membership on the map in order to understand the spatial distribution of each club. By doing so, we are able to explain the connection between associated features of location towards club membership.
Table 1: Empirical Studies of Regional Convergence in Indonesia

<table>
<thead>
<tr>
<th>Study</th>
<th>Period</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garcia and Soelistianingsih (1997)</td>
<td>1975-1993</td>
<td>GDP per capita</td>
</tr>
<tr>
<td>Methods OLS regression on absolute and conditional Beta convergence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding Absolute Convergence ($\beta$) and Conditional convergence ($\beta$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muhammad Firdaus and Zulkornain Yusop (2009)</td>
<td>1983-2003</td>
<td>GDP per capita and some explanatory variables</td>
</tr>
<tr>
<td>Methods OLS, Fixed Effect, Random Effect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding Conditional Convergence ($\beta$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vidyattama (2013)</td>
<td>1999-2008</td>
<td>Williamson Index of GDP per capita and HDI</td>
</tr>
<tr>
<td>Methods Spatial econometrics model (SEM and SAR model)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding Convergence ($\beta$) in HDI, no convergence in GDP per capita</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tirtosuharto (2013)</td>
<td>1996-2012</td>
<td>GDP per capita</td>
</tr>
<tr>
<td>Methods Sigma Convergence, Beta Convergence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finding No $\sigma$ convergence, no $\beta$ convergence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods Club convergence (Phillips and Sul, 2007)</td>
<td></td>
<td>Junior secondary school enrollment</td>
</tr>
<tr>
<td>Finding No overall convergence, 2 clubs convergence</td>
<td></td>
<td>Fertility rate</td>
</tr>
<tr>
<td>Mendez, Carlos (2020)</td>
<td>1990-2010</td>
<td>Three measures of efficiency;</td>
</tr>
<tr>
<td>Methods Classical and Distributional Convergence</td>
<td></td>
<td>Pure, Scala, and Overall efficiency</td>
</tr>
<tr>
<td>Finding Overall convergence, 2 local convergence for pure overall efficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 local convergence for scale efficiency</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors’ documentation.

3.1 Convergence clubs framework

To determine convergence clubs of per capita income among Indonesian districts, we follow the method of club convergence test suggested by Phillips and Sul (2007, 2009). The test itself is based on the decomposition of the panel-data variable of interest in the following way:

$$y_{it} = g_{it} + a_{it}$$

(1)

where $g_{it}$ is a systematic component and $a_{it}$ is a transitory component. Further, to separate common from idiosyncratic components, equation 1 can be transformed with a time-varying factor as follows:

$$y_{it} = \left( \frac{g_{it} + a_{it}}{\mu_t} \right) \mu_t = \delta_{it} \mu_t$$

(2)

where $\delta_{it}$ contains error term and unit-specific component and thus represents idiosyncratic element that varies over time and $\mu_t$ is a common component. To be more specific, the transition path of an observed economy towards its own equilibrium growth path is explained by $\delta_{it}$ while $\mu_t$ depicts a hypothesized equilibrium growth path that is common to all economies. Equation 2 is therefore a dynamic factor model containing a factor loading coefficient $\delta_{it}$ that represents the idiosyncratic distance between a common trending behavior, $\mu_t$, and the dependent variable, $y_{it}$.

Furthermore, to characterize the dynamics of the idiosyncratic component, $\delta_{it}$, Phillips...
and Sul (2007) proposed the following semi-parametric specification:

$$\delta_{it} = \delta_i + \frac{\sigma_i \xi_{it}}{\log(t)} t^\alpha$$

(3)

where \(\delta_i\) represents the heterogeneity of each economy but constant over time, \(\xi_{it}\) is a weakly time dependent process with mean 0 and variance 1 across economies.

Under the condition given in equation 3, convergence occurs when all economies move to the same transition path:

$$\lim_{t \to \infty} \delta_{it} = \delta$$ and \(\alpha \geq 0\)

(4)

In other words, equation 4 is the null hypothesis of convergence.

In order to estimate the transition coefficient \(\delta_{it}\), Phillips and Sul (2007) construct a relative transition parameter, \(h_{it}\), as

$$h_{it} = \frac{y_{it}}{\frac{1}{N} \sum_{i=1}^{N} y_{it}} = \frac{\delta_{it}}{\frac{1}{N} \sum_{i=1}^{N} \delta_{it}}$$

(5)

where common component, \(\mu_t\) in equation 2 is eliminated by dividing the independent variable, \(y_{it}\), with the panel average. Thus, \(h_{it}\) represents the transformation path of economy \(i\) against the level of cross-sectional average, implying the calculation of individual economic behaviors relative to other economies. Then, \(h_{it}\) converges to unity, \(h_{it} \to 1\), when \(t \to \infty\). Thus, the notion of convergence defined in equation 4 can be transformed into following equation that describe the cross-sectional variance of \(h_{it}\),

$$H_t = \frac{1}{N} \sum_{i=1}^{N} (h_{it} - 1)^2 \to 0$$

(6)

where the cross-sectional variance converges to zero, \(H_t \to 0\).

The null hypothesis in equation 4 is verified in counter to the alternative hypothesis \(H_A: \delta_i \neq \delta\) for all \(i\) or \(\alpha \geq 0\). Finally, Phillips and Sul (2007) empirically evaluate this null hypothesis by using the following log t regression model:

$$\log \left( \frac{H_t}{H_{t-1}} \right) - 2 \log \{ \log(t) \} = a + b \log(t) + \epsilon_t$$

for \(t = [rT], [rT] + 1, \ldots, T\) with \(r > 0\)

(7)

where \(rT\) is the initial observation in the regression, which implies that the first fraction of the data (that is, \(r\)) is discarded. Based on Monte Carlo experiments, Phillips and Sul (2007) suggest to set \(r = 0.3\) when the sample is small or moderate \(T \leq 50\). A fairly conventional inferential procedure is also suggested for equation 7. To be more specific, a one-sided t test with heteroskedasticity-autocorrelation consistent (HAC) standard errors is used. In this setting, the null hypothesis of convergence is rejected when \(tb < -1.65\).

### 3.2 Identifying convergence clubs

Note that even though the null hypothesis of overall convergence in full sample is rejected, it does not necessarily mean that the convergence in the sub sample of the panel is not present. Indeed, perhaps the most appealing feature of the model of equation 7 is its ability to reveal the presence of multiple convergence clubs in sub-sample (Phillips and Sul, 2007). In order to do that, Phillips and Sul (2007) suggest an innovative a data-driven algorithm, which can be summarized into following four steps.
1. Ordering: Sample units (districts) are arranged in decreasing order according to their observations in the last period.

2. Constructing the core group: A core group of sample units (districts) is identified based on the maximum $t_k$, which is obtained from a series of sequential estimations of equation 7 for the $k$ largest group ($2 \leq kN$).

3. Deciding club membership: Sample units (districts) not belonging to the core group are reevaluated one at a time with log $t$ regression. A new group is formed when the t-statistic $> 0$.

4. Iteration and stopping rule: The regression model of equation 7 is applied for the remaining sample units (districts). If the process shows the rejection of null hypothesis of convergence, steps 1 to 3 are performed again. The remaining sample units (districts) are labeled as divergent if no core group is found, and the algorithm stops.

3.3 Data

In this study we use annual per capita gross regional domestic product (GRDPcap) in 2010 prices at district level over the period of 2000-2017. We constructed the per capita gross regional domestic product (GRDPcap) by dividing gross regional domestic product (GRDP) with number of populations (Pop) in respective district and time, such as,

$$GRDP_{cap_{it}} = \frac{GRDP_{it}}{Pop_{it}}$$  \hspace{1cm} (8)

Both indicators are collected from Bureau of Statistics. Still, not all of the data are available for every year and lead to missing observations problem caused by splitting up of seven new provinces namely, Bangka Belitung (South Sumatra, 2001), Riau Islands (Riau, 2002), Banten (West Java, 2000), West Sulawesi (South Sulawesi, 2004), Gorontalo (North Sulawesi, 2000), North Maluku (Maluku, 1999), and West Papua (Papua, 2003)\textsuperscript{1}

Since the club convergence test of Phillips and Sul (2007) requires balanced panel data, we solved for missing observations by doing interpolation/imputation. We interpolated the series using a linear regression method with the year and reference districts as candidates of regressors. We followed two rules in interpolation process. First, we selected the reference districts if new districts are established from parts of other districts and second, the year was used as regressor to handle missing observation that is caused by other circumstances. Hence, since we only predicted the missing values from its trend, it would not significantly alters the results of convergence analysis\textsuperscript{2}. Table 2 shows the descriptive statistics of the complete data set. Then, as suggested by previous literature, we transform the GRDP per capita data into log form. Finally, to extract the business cycle component from the logged series, we compute the smoothed GRDP per capita using Hodrick-Prescott (HP) filter with lambda ($\lambda$) parameter being equal to 6,25 as suggested by Ravn and Uhlig (2002).

\textsuperscript{1}Original province and the year of new province establishment are in the parentheses.

\textsuperscript{2}We combined actual data of the new districts and the reference districts, and then compared them to ensure that measurement error caused by the interpolation is minimum. We obtained the result that the new district series follows both its available actual data and its reference district when interpolated (we assume that the series of new districts follow trend of its reference). Details of this interpolation process are provided in Appendix A.
Table 2: Descriptive Statistics of GRDPcap, 2000-2017

<table>
<thead>
<tr>
<th>Club</th>
<th>Observation</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRDPcap</td>
<td>9.252</td>
<td>28.701</td>
<td>46.527</td>
<td>2.004</td>
<td>932.664</td>
</tr>
<tr>
<td>lnHP(GRDPcap)</td>
<td>9.252</td>
<td>9.867975</td>
<td>.7555215</td>
<td>7.58643</td>
<td>13.73496</td>
</tr>
</tbody>
</table>

Source: Authors’ calculation.

4 Results: Are there multiple convergence clubs?

4.1 Discovering convergence clubs

After applying the log t test to per capita incomes across 514 Indonesian districts over the 2000-2017 period for the identification of per capita income convergence clubs, we were able to reject the null hypothesis of overall convergence at the 5% significant level ($\hat{b}$ is significantly $< 0$ with t-statistic $-22.27$). This means the convergence for the total districts is not observed, indicating that the growth process of 514 Indonesian districts from 2000 to 2017 does not show single equilibrium steady state. However, as mentioned before log t test can observe the existence of several convergence clusters when we do not observe overall convergence for all districts. Therefore, we applied the test to investigate convergence clubs and the results are presented in Table 3\(^3\). The $\hat{b}$ column corresponds the fitted coefficients and t-statistic is shown in bracket. We found five significant initial clubs as shown in Table 3. Then we check the possibility whether any of the identified clubs can be merged to form bigger convergence clubs. As shown on the middle panel of Table 3, the results of the merging analysis suggest that there is no possibility to shape large convergence clubs by merging the original clubs ($\hat{b}$ is significantly $< 0$ with t-statistic $< -1.65$). Hence, the five subgroups are confirmed as the final convergence clubs. The first final convergence club consists of 6 districts, second club consists of 126 districts, third club consists of 178 districts, fourth club consists most of members of 181 districts and fifth club consists of 23 districts.

The order of convergence clubs is sorted from the districts with highest to the lowest GRDP per capita i.e. the first club refers to the highest GRDP per capita group, the second club is the lower GRDP per capita group, and the last group displays the lowest GRDP per capita group. These club convergence test implies that generally speaking, for the years of observation (2000-2017), the development of income per capita in 514 Indonesian districts can be grouped into five common trends.

Furthermore, measuring the gap between clubs is also useful in order to understand the income disparities among convergence clubs. For this purpose, we show the mean of per capita income of each club in Table 4. The statistics suggest that the gap of income per capita between clubs are arguably large in particular between Club 1 and Club 2 where the average income per capita of districts in Club 1 is about four times larger than those in Club 2. This implies that Club 2 has very little progress in catching up with Club 1. Table 4 also reflects severe income inequality problems among districts in Indonesia where the average income per capita in the last club is only about 3% of the average in the first club.

\(^3\)See Appendix B for complete members of each club.
Table 3: Results of Club Convergence Test

<table>
<thead>
<tr>
<th>Club</th>
<th>Initial Club [Members]</th>
<th>Merging Test</th>
<th>Final Club [Members]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>514</td>
<td>-0.523 [-22.279]</td>
<td>Club 1 + Club 2 6</td>
</tr>
<tr>
<td>Club 1</td>
<td>6</td>
<td>0.417 [4.968]</td>
<td>Club 3 + Club 4 178</td>
</tr>
<tr>
<td>Club 2</td>
<td>126</td>
<td>-0.084 [-1.521]</td>
<td>Club 2 + Club 3 126</td>
</tr>
<tr>
<td>Club 3</td>
<td>178</td>
<td>0.370 [5.257]</td>
<td>Club 2 + Club 3 126</td>
</tr>
<tr>
<td>Club 4</td>
<td>181</td>
<td>-0.042 [-1.597]</td>
<td>Club 2 + Club 3 126</td>
</tr>
<tr>
<td>Club 5</td>
<td>23</td>
<td>0.486 [6.551]</td>
<td>Club 4 + Club 5 181</td>
</tr>
</tbody>
</table>

Note: Reject the null hypothesis of convergence at the 5% significant level. The numbers in brackets represent the number of districts in each club.

Source: Authors’ calculation.

4.2 Transition path of clubs

In this section, we analyze the convergence process of each member in each club by showing the curve of growth transition path Phillips and Sul (2007). However, it is important to first understand the conceptual framework of transition path curve.

Two lines in Figure 1 represent two different growth transition of economies, or districts in this particular study. District 1 and District 2 depart from different initial level of income. As the development process continues, at one point, the relative growth transition of both districts converge to the same value. More specifically, District A (District B) converge from higher (lower) level of initial income to the average level of cross section income level. From a growth transition path perspective, the transition line of District 2 is more likely represent the typical developing regions as it converges from initial lower level of income to the higher level, also known as "catch-up effect".

Table 4: GDP Per Capita of Clubs, 2000-2017

<table>
<thead>
<tr>
<th>Club</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Club 1</td>
<td>231.289</td>
<td>196.580</td>
<td>7.058</td>
<td>932.664</td>
</tr>
<tr>
<td>Club 2</td>
<td>56.961</td>
<td>58.557</td>
<td>7.718</td>
<td>658.303</td>
</tr>
<tr>
<td>Club 3</td>
<td>20.090</td>
<td>12.178</td>
<td>3.405</td>
<td>304.400</td>
</tr>
<tr>
<td>Club 4</td>
<td>13.469</td>
<td>7.952</td>
<td>4.005</td>
<td>194.717</td>
</tr>
<tr>
<td>Club 5</td>
<td>7.549</td>
<td>5.959</td>
<td>2.004</td>
<td>59.292</td>
</tr>
</tbody>
</table>

Source: Authors’ calculation.
Figure 1: Illustration of relative transition path

Figure 2 shows the transition path of members in each club by comparing income per capita of each district (in logged form) relative to club’s average. All five clubs exhibits different convergence behaviors and transition paths within the club depending on initial conditions and development process of each districts. We also found some asymmetric transition patterns within the club. On one hand, some districts with significantly higher level of income at initial period (particularly in Club 3, Club 4 and Club 5) loss their performance sufficiently large at the end of period and move towards the average level of the club. On the other hand, none of districts with lower level of income at initial period record significant improvement over time.

Figure 2: Relative Transition Path of Clubs
4.3 The clubs in geographical perspective

In addition, we provide a geographical view of club membership in Figure 3. A few regularities are visible in the Indonesian district convergence clubs shown here. First, the province effect notably obvious, that is, districts belonging to the same province tend to be in the same cluster (Barro, 1991; Quah, 1996). This pattern applies almost to all clubs. For example, districts in provinces of East Kalimantan and Riau tend to be grouped in Club 2. Aceh, West Sumatera, West Kalimantan and Central Kalimantan also show similar structure where most of the districts in these provinces are clustered in Club 3, and finally, mostly the districts in Maluku and Nusa Tenggara provinces assemble Group 4 and 5. More surprisingly, districts belonging to the same club tend to be geographically close. To put it another way, the clubs seem to be spatially concentrated. This could be the indication of some spatial agglomeration effects (Martin and Ottaviano, 2001) and driven by factors like spatial externalities or spillovers (Quah, 1996).

Figure 3: Spatial Distribution of Convergence Clubs, 2000-2017

Second, the highest club is dominated by regions with distinct characteristic i.e. big cities or natural resources rich regions like Central Jakarta (the central district of nation’s capital city), Kediri (largest national tobacco producer), Morowali (the location of recently developing nickel based industrial park), Membramo Raya, Mimika and Teluk Bintuni (natural resources districts in coastal area of Papua island, respectively).

Finally, the spatial distribution also shows the prolonged existence of classical regional development problem in Indonesia, that is, the eastern parts of the archipelago are still lagged in development. It can be seen from the membership of the fifth club where out of 23 members, 21 districts are located in eastern part of Indonesia i.e. South Sulawesi, Nusa Tenggara and Papua.
5 Discussion and robustness checks

5.1 Possible policy implications

The difference in the progress of inter-region development is natural. It is related to the potential variation that each region has, both natural resources and geographical location. In addition, variations in the regional ability to manage their resources and potential are also factors that differentiate the success rate of development in each region. Behind the ability of Indonesian economy to maintain robust economic growth after the Asian Financial Crisis in 1997/98, there are still problems between regions gap that could potentially be a problem in the future. Regional income inequality not only could trigger social dispute stemmed from the perception of injustice among fellow communities, but also could pose downside risks to the national economic growth.

To reduce regional income inequality, Indonesian government need to have a clear and accurate picture of regional imbalances among regions. In this context, the results of this study suggests that in general the growth path of income per capita among 514 Indonesian districts during the period of 2000-2017 does not converge to the same steady state level. Similar to Kurniawan et al. (2019), this finding implies the absence of global convergence of income per capita among Indonesian regions. Rather, this study finds that the growth process of Indonesian districts constitutes five local convergence clubs. Income gap among these five clubs is also considerably large, suggesting that the potential regional development policies might be different across clubs. For example, the development policies for districts in Club 1 might be directed to the attempts of seeking for new sources of growth to prevent the drops from the high income level position. Meanwhile, majority of districts in Club 2, 3 and 4 could focus their program on developing the middle sized-cities and more programs on connectivity, infrastructures and basic public services should be implemented in regions in Club 5.

Furthermore, spatial distribution of clubs can provide non-trivia information for inter-provincial policy making in the context of reducing income inequality. For example, Figure 3 shows some districts in Club 5 share the border with districts in Club 1, implying the potential to further strengthen positive spillover from rich regions to poor regions that geographically close. In Papua province for instance, Pegunungan Arfak (Club 5) is the direct neighbor of Teluk Bintuni (Club 1), Deiyai, Puncak, and Nduga (Club 5) share the border with Mimika (Club 1), while Puncak Jaya and Tolikara (Club 5) are the neighbor of Membramo Raya (Club 1). Among others, some possible inter-provincial policies are through providing connectivity and promoting trade between these regions. The same fashion can be also applied in some similar regions such as Aceh Timur in Aceh province, Blora in Central Java province and Jeneponto in South Sulawesi province.

Finally, given the persistence of west-east development gap that is also observed in this study, the central government policies to support the development of physical infrastructures and basic public services provision in eastern part of the archipelago are highly favorable to reduce regional income inequality.

5.2 Robustness checks

5.2.1 Testing the convergence of each club

Now, to evaluate the robustness of the clubs formed by the Phillips and Sul approach Phillips and Sul (2007), we test the existence of convergence among them under the
notion of classical convergence. Further, we also evaluate the convergence speed of each club.

The results of classical convergence test are presented in Table 5. The $\beta$ coefficient of income per capita is negative and statistically significant for all five clubs formed by the convergence club approach, suggesting the evidence of convergence. This also implies that income per capita converge when the districts share common fundamental characteristics. Also, the results presented in Table 5 show that the speeds of convergence for Clubs 1 to 5 are 0.05%, 0.03%, 0.05%, 0.04%, and 0.04%, respectively. Lastly, the fourth column of Table 5 clearly shows that the districts in Club 3 converge at a faster pace than the other clubs. This is also evidently seen from Figure 4, where the distribution of per capita income in Club 3 is the least dispersed compared to other clubs, making the convergence process faster. In summary, the results of classical convergence test support the convergence clubs findings.

<table>
<thead>
<tr>
<th>Club</th>
<th>Beta Coeff</th>
<th>Convergence Speed</th>
<th>Half-life (years)</th>
<th>R-Square</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Club 1</td>
<td>-0.54</td>
<td>0.05</td>
<td>15.09</td>
<td>0.86</td>
<td>0.00</td>
</tr>
<tr>
<td>Club 2</td>
<td>-0.43</td>
<td>0.03</td>
<td>21.14</td>
<td>0.74</td>
<td>0.00</td>
</tr>
<tr>
<td>Club 3</td>
<td>-0.59</td>
<td>0.05</td>
<td>12.98</td>
<td>0.83</td>
<td>0.00</td>
</tr>
<tr>
<td>Club 4</td>
<td>-0.51</td>
<td>0.04</td>
<td>16.36</td>
<td>0.69</td>
<td>0.00</td>
</tr>
<tr>
<td>Club 5</td>
<td>-0.52</td>
<td>0.04</td>
<td>15.05</td>
<td>0.64</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Source: Authors’ calculation.

5.2.2 Applying different smoothing parameter (lambda) on HP filter

We also test the convergence club method by using smoothing parameter ($\lambda$) parameter of Hodrick-Prescott (HP) filter equal to 400 which is used in the study of Phillips and
Similar with the results discussed in previous section, the existence of global convergence is rejected at the 5% significant level ($\hat{b}$ is significantly $< 0$ with t-statistic $-23.02$). Then we proceeded with the convergence clubs test by following the same procedures discussed in Section 2. We found initial twelve convergence clubs. Next, we apply the merging test procedure to investigate whether any of the initial subgroups can be merged to form convergence clubs with larger number of members.

As the results, we also found five final significant convergence clubs. The original first and second convergence clubs are merged into first convergence club with 14 members while third original club become second club with 106 members. Next, the third new club is formed by fourth, fifth and sixth original clubs with 240 members and becomes the largest final club (47 percent of total number of districts). Next, the fourth final club is constructed by club 7, 8, 9, 10 and 11 in the original club with 132 members. Finally, the twelfth (the last) original club stays unmerged with 22 members (details of results are presented in Appendix C). This club merging process shows interesting pattern i.e. most of the clubs in the middle range of initial twelve groups can be merged, while one lowest club can not be merged. Also, based on the statistics shown in Table 6, the mean column shows that the biggest gap is between Club 1 - Club 2 (around 4 times), while the gap between Club 3 - Club 4 and Club 4 - Club 5 are less than two times.

<table>
<thead>
<tr>
<th>Club</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Club 1</td>
<td>197.173</td>
<td>163.309</td>
<td>6.900</td>
<td>922.607</td>
</tr>
<tr>
<td>Club 2</td>
<td>50.491</td>
<td>38.901</td>
<td>4.576</td>
<td>275.009</td>
</tr>
<tr>
<td>Club 3</td>
<td>20.509</td>
<td>12.286</td>
<td>2.822</td>
<td>299.126</td>
</tr>
<tr>
<td>Club 4</td>
<td>11.801</td>
<td>3.198</td>
<td>4.234</td>
<td>31.113</td>
</tr>
<tr>
<td>Club 5</td>
<td>5.804</td>
<td>1.730</td>
<td>1.971</td>
<td>10.381</td>
</tr>
</tbody>
</table>

Table 6: GDP Per Capita of Each Final Club, 2000-2017 (with lambda = 400)

Source: Authors’ calculation.

6 Conclusions

The development process in many countries is mainly aimed to achieve the ultimate goal, namely welfare for all society. One indicator of the of is equality in per capita income level, both between the class community and between regions. In the same context, this study documents the evolution of disparities in income per capita among Indonesian districts in post decentralization era. In particular, using a complete district level data of per capita income in 514 Indonesian districts from 2000 to 2017, the convergence test proposed by Phillips and Sul (2007) is applied to evaluate whether all districts are converging to a common steady-state path.

The findings are documented as follows. First, there is no overall convergence in income per capita among Indonesian districts. Rather, we found five convergence clubs that describe the evolution of income growth among Indonesian districts. Second, as we observed a large income gap between clubs, the "catch-up effects" seems to be exist within the club but not between clubs. This prevalent pattern calls for differentiated regional development policies across clusters. Third, although in general districts belonging to the same province tend to cluster together (called as province effect by Barro,
some rich districts are geographically close with poor districts in the same province. As clearly described in the discussion section, province like Papua, Aceh, Central Java and South Sulawesi might consider about inter-provincial development policies to reduce income gap between its districts by fostering spillovers from rich to poor regions.

Finally, the spatial distribution of convergence clubs clearly shows the persistence of west-east development gap. Thus, central government should coordinate regional policies to support the development in eastern part of Indonesia to close the gap with the west. Among other priorities, the provision of physical infrastructures and basic public services in eastern part of the archipelago is highly favorable.
Declaration

1. Ethics approval and consent to participate.
   Not applicable.

2. Consent for publication.
   Not applicable.

3. Availability of data and materials.
   The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

4. Competing interests.
   The authors declare that they have no competing interests.

5. Funding.
   The authors received no financial support for the research, authorship, and/or publication of this article.

6. Authors’ contributions.
   All authors developed the conceptual and theoretical framework. Both first and second author performed the analytic calculations and performed the numerical simulations. All authors contributed to the final version of the manuscript. Third author supervised the project.

7. Acknowledgements.
   Not applicable.
References


Appendix A

Districts Interpolation

The Indonesia Central Bureau of Statistics has published periodically the data of per capita Gross Regional Domestic Product (per Capita GRDP) at district (regencies and cities) level at current price. However, the sequential time series database of per Capita GRDP 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 A). 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.

![Figure A: Number of Districts in Indonesia since 2000](image)

In this paper, we construct GRDP per capita at constant price by calculating GRDP at constant price and population at all districts. To do so, we conducting interpolation or imputation for all missing value in both the GRDP 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 ?. 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 Regency 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.

.1 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 Panel A1 (in Figure B), 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.

.2 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 Panel A2 (in Figure B). The comparison of imputed GDP data for Kerinci Regency and Sungai Penuh City can be seen in Panel A3 (in Figure B) and the sum of both districts is equal with the reference districts.

.3 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.

.4 Calculating Per Capita GDP Data

Since the interpolation for both GRDP and total population at district level are done, the next step is calculating per capita GRDP at constant price for all districts. Per capita GRDP at constant price for districts level are obtained by dividing GRDP (as numerator) by the total population (as de-numerator) in each district. Both database for GRDP and per capita GRDP at districts level in this paper have been adjusted using the latest constant price 2010. As an example, the comparison of per capita GRDP at
districts level between Kerinci Regency and Sungai Penuh City can be seen in Panel A4 (in Figure B).

Figure B: 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.

Appendix B

Table of Club Membership
Table A: List of Districts in Each Club

<table>
<thead>
<tr>
<th>Club</th>
<th>Districts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Club 1 (6)</td>
<td>Jakarta Pusat, Kediri, Mamberamo Raya, Mimika, Morowali, Teluk Bintuni.</td>
</tr>
</tbody>
</table>

21
<table>
<thead>
<tr>
<th>Club</th>
<th>List of Districts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Club 3 (178)</td>
<td>Pulang Pisau Raja, Raja Ampat, Lebong Rokan, Hulu Sambas, Sambas, Samosir, Sarolangun, Serang, Serdang Bedagai, Kei, Tagulandang Biaro, Sidenreng Rappang, Sigi, Sijunjun, Simalungun, Sinjai, Sleman, Solok, Soppeng, Sragen, Sukamara, Sukoharjo, Sumbawa, Supiori, Tabanan, Takalar, Tana Toraja, Tanah Bumbu, Tanah Datar, Tanah Laut, Tapin, Tebo, Teluk Wondama, Toba, Samosir, Tojo Una-Una, Toli-Toli, Toraja Utara, Tuban, Tulang Bawang, Tulang Bawang Barat, Tulungagung, Yalimo</td>
</tr>
<tr>
<td>Club 5 (23)</td>
<td>Aceh, Timur, Blora, Buru, Deiyai, Jeneponto, Bima, Kupang, Lanny, Jaya, Lambata, Manggarai, Manggarai, Barat, Manggarai, Timur, Nagekeo, Nduga, Pegunungan, Arfak, Puncak, Puncak, Jaya, Rote, Ndau, Sabu, Raijua, Sumba, Barat, Daya, Sumba, Tengah, Tolikara, Yahukimo</td>
</tr>
</tbody>
</table>
### Appendix C

Club Merging with lambda = 400

Table B: Results of Club Convergence Test (with lambda = 400)

<table>
<thead>
<tr>
<th>Initial Club</th>
<th>Club Merging Test</th>
<th>Final Club</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample [514]</td>
<td>-0.530 [-23.017]</td>
<td>Club 1 [14] 0.302 [3.787]</td>
</tr>
<tr>
<td>Club 2 [10]</td>
<td>0.271 [3.252]</td>
<td>Club 2 + Club 3 -0.072 [-1.494]</td>
</tr>
<tr>
<td>Club 3 [106]</td>
<td>0.022 [0.423]</td>
<td>Club 3 + Club 4 -0.237 [-6.390]</td>
</tr>
<tr>
<td>Club 4 [186]</td>
<td>0.250 [3.888]</td>
<td>Club 4 + Club 5 0.019 [0.583]</td>
</tr>
<tr>
<td>Club 5 [50]</td>
<td>0.026 [1.254]</td>
<td>Club 5 + Club 6 0.084 [3.848]</td>
</tr>
<tr>
<td>Club 12 [22]</td>
<td>-0.002 [-0.069]</td>
<td>Club 5 [22] -0.002 [-0.069]</td>
</tr>
</tbody>
</table>

Note: Reject the null hypothesis of convergence at the 5% significant level.

The numbers in brackets represent the number of districts in each club.

Source: Authors’ calculation.
Figures

Figure 1
Illustration of relative transition path

Figure 2
Relative Transition Path of Clubs

Figure 3
Spatial Distribution of Convergence Clubs, 2000-2017

Figure 4
Beta Convergence of Clubs
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix.pdf