Environmental Efficiency and Regional Convergence Clusters in Japan: A Nonparametric Density Approach

Mendez-Guerra, Carlos

Nagoya University

18 February 2019

Environmental Efficiency and Regional Convergence Clusters in Japan: A Nonparametric Density Approach

Carlos Mendez-Guerra*
Nagoya University
Graduate School of International Development

(Preliminary Version 1.0)

Abstract

This paper studies environmental efficiency convergence across the prefectures of Japan over the 1992-2008 period. Using a novel nonparametric density estimation clustering framework, two alternative indicators of environmental efficiency are contrasted: a conventional indicator, based on the ratio of gross regional product to CO2 emissions, and a more comprehensive indicator, based on the data envelopment analysis (DEA) model. Results show, on the one hand, a lack of intra-distributional mobility and potentially a unique convergence cluster when using the more conventional indicator. On the other hand, large backward mobility and at least two convergence clusters are identified when using the DEA-based indicator of environmental efficiency. The paper concludes arguing the importance of accounting for production inputs, as they appear to be driving the formation of regional convergence clusters in Japan.

Keywords:

environmental efficiency, convergence, nonparametric density, Japan

---

* I am grateful to all participants of the Economic Development Policy and Management seminar at Nagoya University for the insightful comments and suggestions.
Contact Email: `carlos@gsid.nagoya-u.ac.jp`
Homepage: `https://carlos-mendez.rbind.io`

1 Introduction

In accordance with the Paris agreement, by the year 2030, Japan is committed to reduce its CO2 emissions by 26 percent relative to its emissions in 2013. To achieve this objective, Japan aims not only to increase the national—average—level of environmental efficiency, but also aims to reduce its cross-regional dispersion (Ministry of Environment, 2013). Although considerable progress has been made on the former, reducing regional gaps in environmental efficiency remains challenging. In this context, this paper studies the evolution of environmental efficiency dispersion across 46 prefectures of Japan over the 1992-2008 period.

In particular, by using both classical measures of regional convergence and a novel framework of distributional convergence, two alternative measures of environmental efficiency are contrasted. The first measure is a conventional environmental efficiency indicator (CEEI for short) that focuses only on the desirable and undesirable outputs of the production process. It is constructed as the ratio of gross regional product to CO2 emissions. The second measure (DEEI for short) is based on the data envelopment analysis (DEA) model. It includes not only desirable and undesirable outputs, but also information about production inputs such as the weight of buildings and roads, the monetary value of private capital stock, and the number of employees.

In a recent work, Eguchi (2017) compares regional performance in these two indicators. His work provides an informative and detailed ranking of Japanese prefectures for each environmental efficiency indicator in the year 1992 and 2008.[1] Although, from this ranking, it is possible to identify the—average—regional progress in CEEI and the—average—regional regress in DEEI, a formal analysis of regional dispersion and convergence is still needed for a more complete understanding of the regional environmental efficiency dynamics of Japan.

The empirical literature on economic convergence that started in the late 1980s has suggested a variety of statistical frameworks for evaluating the dynamics of inequality. The survey article of Magrini (2004) attempts to sum-

[1] Based on this table, the author points out some ranking differences between the indicators in each year.

marize some of these frameworks in the context of the regional territories of a country. On the one hand, there are the classical convergence frameworks that are based on single summary measures. Among them, the most commonly used are the beta and sigma convergence statistics.[2] On the other hand, there are the distributional convergence frameworks that are based on modeling the shape and internal dynamics of the entire regional distribution.[3]

By applying these two convergence frameworks, and other related analyses, this paper extends the findings of Eguchi (2017) in six fronts. First, it documents a U-shaped relationship between CEEI and DEEI. Second, it points out that low-efficiency regions are not catching-up with the high-efficiency regions (i.e., lack of beta convergence). Third, across multiple indicators of dispersion, environmental differences across prefectures have actually increased (i.e., lack of sigma convergence). Forth, the nonparametric distribution of CEEI shows a constant spread with a unique mode, while the distribution of DEEI shows an increasing spread from its lower tail and a drastic transformation of its multiple modes. Fifth, the intra-distribution dynamics of CEEI mostly suggest regional stagnation, while the dynamics of the DEEI mostly suggest backward mobility within the distribution.

The last finding is based on the novel clustering approach via nonparametric (unconditional) density estimation of Azzalini et. al (2007) and Menardi and Azzalini (2014). In the case of DEEI, this framework suggests the existence of tree groups of regions: a "Tokio-led" convergence cluster, a "median-led" convergence cluster, and a group of transitioning regions. Moreover, the conditional density estimation approach of Hyndman et. al (1996) helps us extend this finding and points out that not all transitioning regions may necessarily end up in any of the two superior clusters. Conditional on their initial position, it is possible that some transitioning regions could end up converging to two additional bottom clusters, one of which is considerably below the median.

The rest of this paper is organized as follows. Section 2 describes the empirical framework, both in terms of the classical convergence approach and the distributional convergence approach. Section 3 describes the data and presents

[2]See Sala-i-Martin (1996) for an overview of this classical approach to convergence analysis.
[3]See Magrini (2009) for an overview of this distributional approach to convergence analysis.

the results of the classical convergence approach. Section 4 presents the main results of the distributional convergence approach. Finally, Section 5 offers some concluding remarks.

2 Methods

2.1 Classical Convergence Approach

The seminal work of Baumol (1986) launched a large number of studies that aim to test the convergence hypothesis across countries and regions within a country. Among the early studies, the work of Barro and Sala-i-Martin (1992), in the cross-country context, and Barro and Sala-i-Martin(1992), in the context of the prefectures of Japan and the states of the US, use the so-called beta convergence approach. In their framework, economic growth is a function of the initial level of income. If, in a regression setting, a statistically significant negative coefficient is found, then poor economies tend to grow faster than rich economies, and consequently a process of convergence takes place. More specifically, Barro and Sala-i-Martin (1991) suggest the following growth equation:

$$\frac{1}{t}\, log\left(\frac{y_t}{y_0}\right) = \gamma - \frac{\left(1 - e^{-\beta t}\right)}{t}\, log\, y_0 + u_t, \tag{1}$$

where the left side is the average growth rate of the variable y, γ is a constant, t is a time index, β is the speed of convergence, y_0 is the initial level of the variable under study, and u_t is a random (white noise) disturbance.[4] In the cross-country context, the parameter γ should be replaced by a set of variables that act as proxies for the unobservable technological and institutional differences that exist across countries. Within countries, however, these differences are relatively minor, and thus γ is commonly assumed constant across regions.

In addition to the rate of convergence (β), a second parameter of interest, known as the "half-life" measure of convergence, can also be computed as:

$$half\text{-}life = \frac{log2}{\beta}. \tag{2}$$

[4]Note that for the purposes of the present paper, the variable y represents the level of environmental efficiency.

This second parameter is particularly informative because it measures the time that a representative economy (or region) needs to halve the gap between its initial level and its convergence equilibrium.

A more general notion of convergence has to do with the reduction of the cross-sectional dispersion of a variable over time. In the convergence literature, this is the so-called sigma convergence approach, and it is commonly measured by the standard deviation of the natural logarithm of the variable y or by its coefficient of variation (CV). Specifically, a decrease in the sigma convergence statistic indicates that the cross-sectional units of the sample (e.g., the regions of Japan) are becoming increasingly more similar in terms of the level of the variable y.

Regarding the relationship between beta and sigma convergence, Quah (1993) and Sala-i-Martin (1996), show that beta convergence is a necessary—but not sufficient—condition to achieve sigma convergence. To derive this relationship, let us first restate Equation 1 in a two-period setting:

$$log\left(\frac{y_t}{y_{t-1}}\right) = \gamma - \left(1 - e^{-\beta}\right)\, log\, y_{t-1} + u_t. \tag{3}$$

Next, apply the variance operator (σ_t^2) to the variable $log\, y_t$ in Equation 3 and assume the white noise properties of u_t. After some algebraic operations, we can find that σ_t^2 evolves over time according to the following equation:

$$\sigma_t^2 = e^{-2\beta}\sigma_{t-1}^2 + \sigma_{ut}^2. \tag{4}$$

Equation 4 is just a first-order differential equation problem. If we assume that the variance of the random disturbance is constant over time ($\sigma_{ut}^2 = \sigma_u^2$), then the solution to this problem is

$$\sigma_t^2 = \frac{\sigma_u^2}{1 - e^{-2\beta}} + \left(\sigma_0^2 - \frac{\sigma_u^2}{1 - e^{-2\beta}}\right) e^{-2\beta t}. \tag{5}$$

Furthermore, in a the steady-state equilibrium $\left(\sigma_t^2 = \sigma_{t-1}^2 = \sigma^2\right)$, the cross-sectional dispersion of y only depends on the beta convergence parameter (β)

and the dispersion of the random disturbance (σ_u^2). That is,

$$\sigma^2 = \frac{\sigma_u^2}{1 - e^{-2\beta}}. \tag{6}$$

From Equation 6, it is clear that beta convergence (an increase in β) is not a sufficient condition for sigma convergence (a decline in σ^2). The effect of beta could be attenuated or even reversed by an increase in the dispersion of random disturbances (σ_u^2), which are commonly interpreted as random shocks to the economies of this dynamical system.

2.2 Distributional Convergence Approach

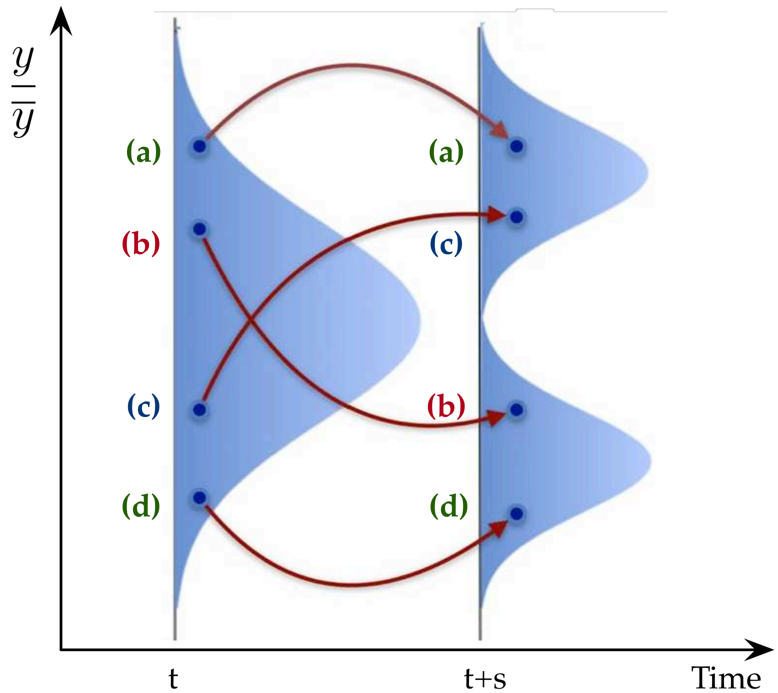
One of the common criticisms to the classical measures of convergence (i.e., beta and sigma) is that they fail to provide information about nonlinear dynamics, heterogenous patterns of convergence, and the formation of local convergence clusters. In essence, beta and sigma convergence are just summary measures of a representative region (i.e., the average region) and, as such, they abstract from important patterns of cross-sectional heterogeneity. Motivated by this limitation, Quah (1993; 1996; 1997) introduced the distributional convergence approach, which goes beyond the performance of the representative region and evaluates the dynamics of the entire cross-sectional distribution.[5]

Figure 1 presents an intuitive graphical summary of the distributional convergence approach. First, in this example, the dynamics of the external shape of the distribution point out not only the lack of overall convergence, but also the formation of two local convergence clusters. Second, the intra-distributional dynamics emphasize the three sources of the clustering process. Some regions, such as (a) and (d), could show large and persistent differences over time. Others, such as (c), could move forward and catch-up with the upper cluster. Lastly, the remaining regions, such as (b), could move backward and join the bottom cluster.

More formally, let us first economize notation and define the relative

[5] See Epstein et al. (2003), Magrini (2004, 2009), Bianco (2016) or Mendez-Guerra (2017; 2018) for more recent and comprehensive presentations of this framework that include not only an evaluation of the intra-distributional dynamics, but also the computation of the long-run (ergodic) equilibrium of the distribution.

***Figure 1.* External-shape dynamics and intra-distributional dynamics**



Note: $y/\bar{y}$ indicates that the value of the variable y is expressed in relative terms. $\bar{y}$ indicates the convergence benchmark, which is usually the cross-sectional mean, median, or any other referential observation.
Source: Adapted from Quah (1993).

level[6] of the variable y at time t as $x \equiv \frac{y_t}{\bar{y}_t}$, and at time $t+s$ as $z \equiv \frac{y_{t+s}}{\bar{y}_{t+s}}$. Next, let us define $f_t(x)$ as the entire cross-sectional distribution at time t and $f_{t+s}(z)$ is the distribution at time $t+s$. In the distributional convergence literature, the evolution from time t to time $t+s$ is commonly modeled as a first-order autoregressive process of a time-homogeneous Markov chain. That is,

$$\underbrace{f_{t+s}(z)}_{Future\ Distribution} = \int \underbrace{f_{t+s|Z_t=x}(z)}_{Transitional\ Operator} \underbrace{f_t(x)}_{Initial\ Distribution} dx, \tag{7}$$

where the transformation of the initial distribution, $f_t(x)$, into the future distribution, $f_{t+s}(z)$, is mapped out by the transitional dynamics summarized by the probability operator, $f_{t+s|Z_t=x}(z)$, which is commonly referred in the literature

[6]For this analysis, the environmental efficiency level of each region is expressed relative to that of Tokio.

as the stochastic kernel.

To estimate the transitional operator, most studies use nonparametric methods (Wand and Jones, 1995; Henderson and Parmeter, 2015). In particular, it is common to define this operator as a conditional distribution:

$$f_{t+s|Z_t=x}(z) = \frac{f_{t,t+s}(z,x)}{f_t(x)}, \tag{8}$$

where $f_{t,t+s}(z,x)$ is an unconditional joint distribution and $f_t(x)$ is the marginal distribution of x. Next, let us define this unconditional joint distribution as:

$$f_{t,t+s}(z,x) = \frac{1}{nh_z h_x} \sum_{i=1}^{n} K_z\left(\frac{z-z_i}{h_z}\right) K_x\left(\frac{x-x_i}{h_x}\right), \tag{9}$$

where K_z and K_x denote Gaussian kernel functions, and h_z and h_x denote the smoothing parameters (flexible bandwidths) for z and x respectively.[7] Lastly, similar to the joint distribution, the marginal distributions, $f_t(x)$, is estimated using a single Gaussian kernel function with a flexible bandwidth.

At this point, it is important to emphasize that intra-distributional dynamics[8] are driven by two components: the unconditional joint distribution, $f_{t,t+s}(z,x)$, and the initial marginal distribution, $f_t(x)$. To study the effects of the former, I implement the approach of Azzalini et. al (2007) and Menardi and Azzalini (2014). To study the effects of the latter, I implement the approach of Hyndman et. al (1996).

The work of Azzalini et. al (2007) and Menardi and Azzalini (2014) is useful not only to estimate the unconditional joint distribution distribution, $f_{t,t+s}(z,x)$, but also to identify and distinguish between convergence clusters and transitioning regions. In a series of papers, these authors have develop an innovative clustering algorithm that is based on the local modes of an estimated nonparametric distribution. More specifically, they exploit recent advances in computational geometry to identify spatially connected regions.

The work of Hyndman et al. (1996) is useful to evaluate the complete process of transitional dynamics, $\frac{f_{t,t+s}(z,x)}{f_t(x)}$. Compared to the original conditional

[7]Optimal smoothing parameters are derived from the the minimization of the asymptotic mean integrated square error (AMISE).

[8]Note that the terms intra-distributional dynamics, transitional operator, stochastic kernel and conditional distribution ultimately refer to the same concept. That is, the movement of the regional units within the cross-regional distribution.

density estimator of Quah (1993,1997), the estimator of Hyndman et al. (1996) has better asymptotic mean bias properties (Arbia et. al,2006). These authors also propose two novel graphical methods for visualizing the estimated conditional densities. The first one is somewhat similar to Figure 1. It is stacked conditional density (SCD) plot that displays a sequence of densities side by side in a perspective fashion. This new graphical device is conceptually similar to a Markovian transition matrix. That is, for a given level of a variable at time t, the SCD traces its evolution and illustrates its final configuration at the time $t+s$. The second graphical method is called high density region (HDR) plot. Along a series of vertical strips, the HRD plot illustrates the projection on the xz plane of the conditional density of z on x. Darker-shaded stripes indicate higher density areas, and the mode for each conditional density is shown as bullet point.

3 Data and Some Stylized Facts

3.1 Data and Sample

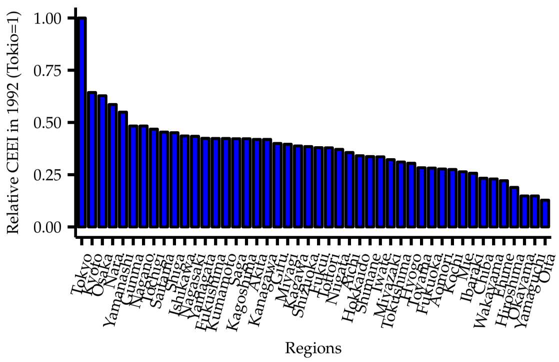
This paper studies dispersion and convergence patterns of environmental efficiency across 46 prefectures of Japan[9] in the year 1992 and 2008. The dataset for this analysis is from Eguchi (2017), who computes both a conventional environmental efficiency indicator (CEEI for short), based on the ratio of gross regional product to CO2 emissions, and a more comprehensive indicator of environmental efficiency (DEEI for short), based on the data envelopment analysis (DEA) model. As noted by this author, the main advantage of the DEA model is that it can include not only desirable and undesirable outputs (i.e., gross regional product and CO2 emissions), but also production inputs such as buildings, roads, private capital stock, and the number of employees.[10]

Figure 2 shows a comparative ranking between the conventional environ-

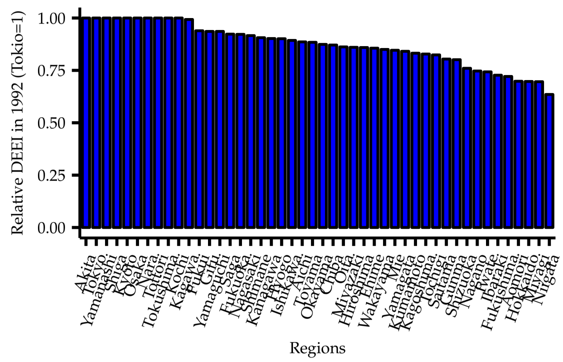
[9]Due to lack of systematic data, the prefecture of Okinawa is excluded from the analysis.

[10]To estimate the relative efficiency index in the DEA framework, Eguchi (2017) uses both physical and monetary capital inputs. To measure the former, estimates of the weight of buildings and roads are taken from Tanikawa et al. (2015). To measure the latter, estimates are taken from the R-JIP database of the Research Institute of Economy, Trade and Industry of Japan. Data on the number of employees and the gross regional product are also from this original source.

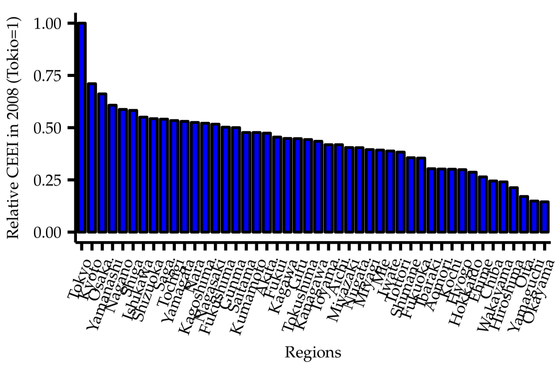
***Figure 2*. Comparison between the conventional environmental efficiency indicator (CEEI) an the DEA-based environmental efficiency indicator (DEEI)**



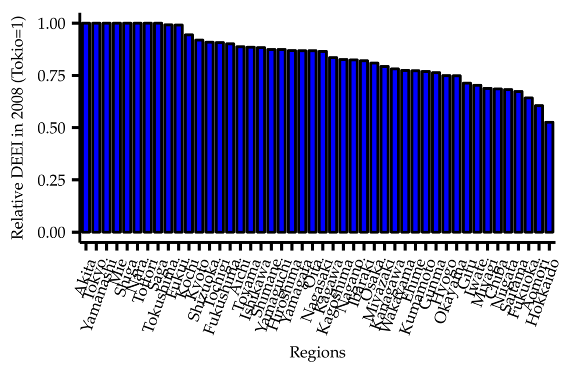
(a) CEEI ranking in 1992



(b) DEEI ranking in 1992



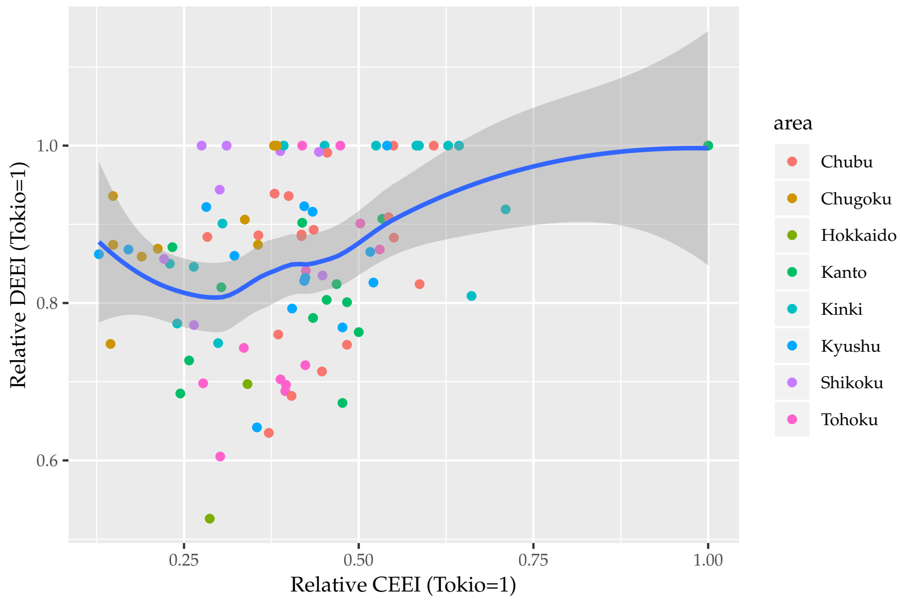
(c) CEEI ranking in 2008



(d) DEEI ranking in 2008

Note: To facilitate the comparison, both indicators have been rescaled relative to the efficiency level of Tokio. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency. In this article, the DEA indicator is based on measures of gross regional production, CO2 emissions, and production inputs (buildings, roads, private capital stock, and the number employees).
Source: Eguchi (2017, p.12), Table 3.

Figure 3. A U-shaped relationship between CEEI and DEEI?



Note: CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency. In this article, the DEA indicator is based on measures of gross regional production, CO2 emissions, and production inputs (buildings, roads, private capital stock, and the number employees).
Source: Author's calculations using data from Eguchi (2017).

mental efficiency indicator (CEEI) an the DEA-based environmental efficiency indicator (DEEI). The first notable difference is that most CEE values tend to be below 0.50, while most DEE values tend to be above 0.50. To some extend, this difference is expected given the construction of both indicators. For instance, the CEEI is an unbounded absolute indicator, which means that the ratio of gross regional product to CO2 emissions can continuously grow over time and it is not necessarily bounded within a certain range of values. The DEEI, however, is a bounded relative indicator, which means that it necessarily takes a value within a certain range. In the context of the DEA framework, this efficiency value is between 0 and 1.[11]

When studying the empirical relationship between both indicators, Figure

[11]In the original results of Eguchi (2017), the direction of efficiency improvements is from 1 to 0. In other words, maximum efficiency is attained when the DEEI equals 0. In the present study, however, the direction of efficiency improvements is reversed, meaning that maximum efficiency is attained when DEEI equals 1. As a result of this adjustment, the comparison between CEEI and DEEI becomes more immediate and intuitive.

3 suggests that the relationship is not only highly non-linear, but also shows signs of a U-shaped pattern. In particular, this U-shaped relationship holds very well when CEEI is between 0 and 0.50, which is the range in which most prefectures are located (See Figure 2). When trying to interpret this pattern, the initial negative relationship between CEEI and DEEI could suggest that initial improvements in environmental efficiency (higher CEEI) are associated with relatively inefficient use of production inputs (lower DEEI). On the other hand, as regions keep developing and improving their environmental performance (higher CEEI), they may also learn how to use their production inputs more efficiently, which improves even further their environmental performance (higher DEEI).

3.2 Lack of Beta and Sigma Convergence

Are the regions with the lowest environmental efficiency scores catching up with the most efficient regions? The results of Table 1a suggest that this is not the case for both indicators of environmental efficiency. Although the beta coefficient is negative, which is a sign of convergence, its p-value is indicates that this coefficient is not statistically significant. Moreover, even if we were to ignore the statistical significance of the coefficient, the speed of convergence is extremely slow. For instance, in the case of the CEEI, its half-life time value suggests that it would take on average 281 years for the current disparities to be halved.

Although—on average—there has been progress in both indicators of environmental efficiency (Eguchi, 2017), the standard sigma convergence analysis suggests that, after more than 15 years, there has been no significant progress in terms of the reduction of environmental efficiency disparities. Table 1 shows that through the lens of two commonly used indicators of dispersion, the standard deviation and the coefficient of variation, the environmental efficiency differences across prefectures have actually increased in most cases. The only exception is the coefficient of variation of the CEEI, which indicates a minor reduction from 0.39 in the year 1992 to 0.36 in the year 2008.

In the convergence econometrics literature, a ratio of initial dispersion over final dispersion larger than one $\left(\frac{\sigma_{1992}}{\sigma_{2008}} > 1\right)$ is typically suggestive of re-

Table 1
Lack of beta and sigma convergence across Japanese prefectures

(a) *Lack of beta convergence*

	Dependent Variable	
	Growth of CEEI (1992-2008)	Growth of DEEI (1992-2008)
Beta coefficient	- 0.039	- 0.214
Speed of convergence	0.002	0.015
Half-life time (in years)	281	46
R-squared	0.02	0.04
P-value of beta	0.41	0.16

(b) *Lack of sigma convergence*

	CEEI		DEEI	
	Standard Deviation	Coefficient of Variation	Standard Deviation	Coefficient of Variation
Dispersion in1992	0.39	0.39	0.12	0.12
Dispersion in 2008	0.40	0.36	0.15	0.14
Dispersion Ratio $\left(\frac{\sigma_{1992}}{\sigma_{2008}}\right)$	0.99	1.08	0.79	0.80
F-Statistic	0.98	0.92	0.62	0.69
P-value	0.95	0.77	0.11	0.21

Note: For the beta convergence analysis, all regressions include a constant term, which is not presented in the table. For the sigma convergence analysis, following the convention of the literature, the standard deviation of the natural logarithm of each indicator is computed. Also following this literature, the coefficient of variation is computed, but this time without its natural logarithm. CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency. In this article, the DEA indicator is based on measures of gross regional production, CO2 emissions, and production inputs (buildings, roads, private capital stock, and the number employees).
Source: Author's calculations using data from Eguchi (2017).

gional convergence. In this context, the results of Table 1b suggest the opposite: the hypothesis of sigma convergence is actually *rejected* in all cases. Even in the previously mentioned exceptional case, the F-statistic and p-value suggest that the reduction in the dispersion is not statistically significant.

4 Distributional Convergence Results

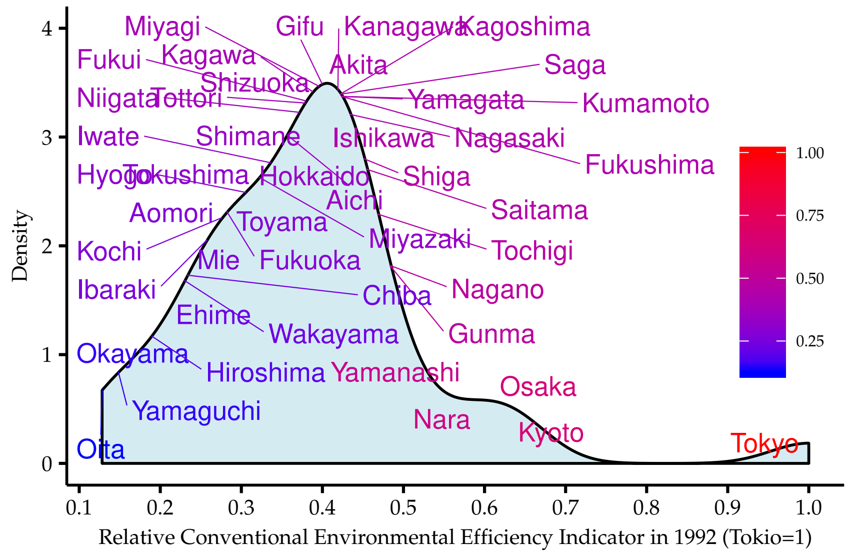
4.1 External-Shape Dynamics

Figure 4 shows the regional composition and dynamics of the nonparametric kernel distribution of CEEI. In the year 1992, we can observe a highly asymmetric environmental efficiency distribution that is characterized by one major density peak at around 41 percent of the environmental efficiency level of Tokio. Also, there is a minor density bump in the range between 50 to 70 percent, which is composed by tree prefectures: Nara, Osaka, and Kyoto. In contrast, in the year 2008, we can observe a relatively more symmetric distribution with only one major peak at around 46 percent of Tokio. Moreover, there is no any minor bump, which indicates that over time Osaka and Nara in particular have been gravitating towards the unique basin of attraction of the distribution.

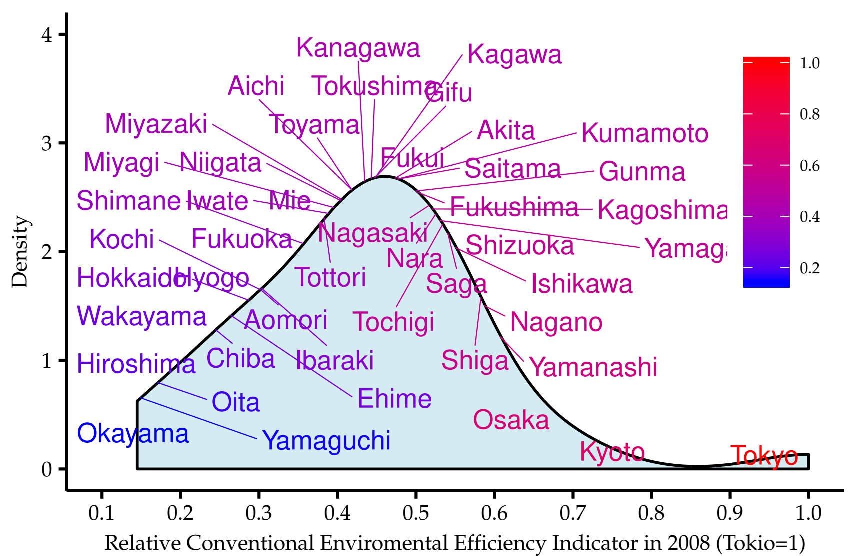
Figure 5 shows the regional composition and dynamics of the nonparametric kernel distribution of DEEI. In the year 1992, we can clearly observe a trimodal distribution. The major peak is located at around 87 percent of the environmental efficiency level of Tokio. In the year 2008, the distribution drastically changed its shape. Starting from the central mode, a considerably large number of regions appear to be moving backwards and, thus, the separation between the lower and central mode has tended to disappear over time. Likewise, the separation between the upper and central mode has tended to disappear, albeit to a lesser extend.

Figure 6 summarizes the distributional changes over time and contrasts both indicators of environmental efficiency. In the case of the CEEI, although the spread of the distribution seems be relatively constant over time (as pointed out in the sigma convergence analysis), the mode of the distribution slightly shifts to the right and becomes more symmetric. The spread of the DEEI, how-

Figure 4. The nonparametric distribution of CEEI: Composition and shape dynamics



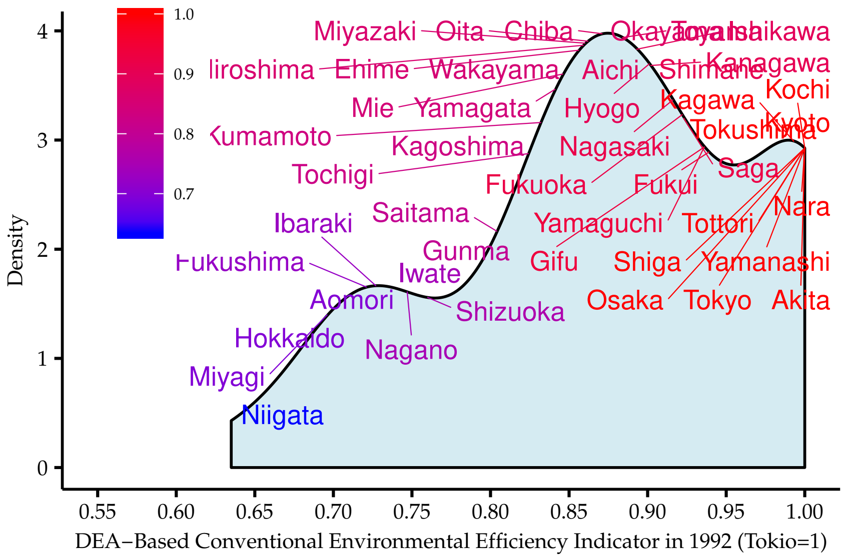
(a) Cross-sectional kernel distribution in the year 1992



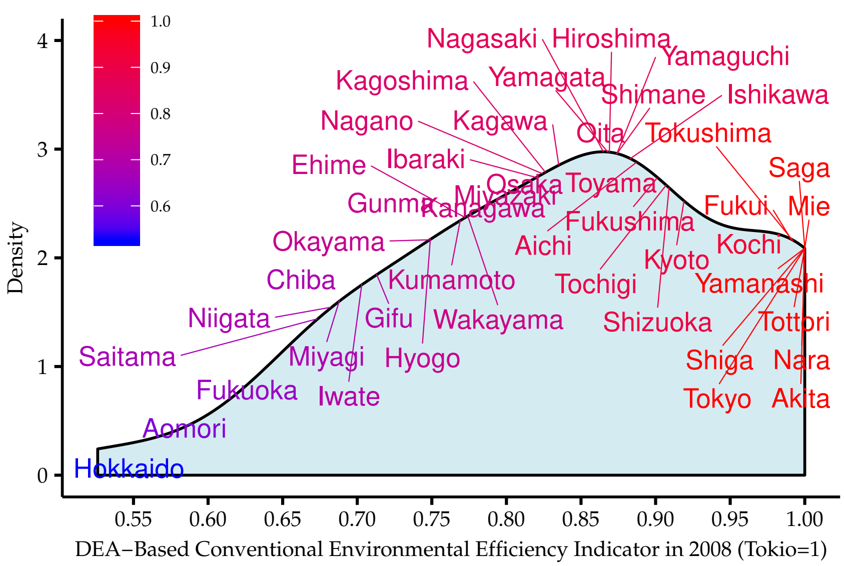
(b) Cross-sectional kernel distribution in the year 2008

Note: CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency.
Source: Author's calculations using data from Eguchi (2017).

Figure 5. The nonparametric distribution of DEEI: Composition and shape dynamics



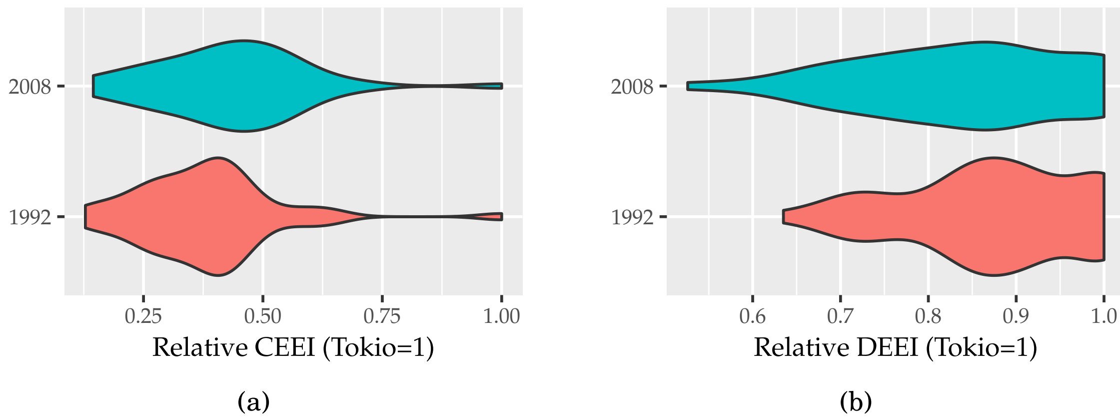
(a) Cross-sectional kernel distribution in the year 1992



(b) Cross-sectional kernel distribution in the year 2008

Note: CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO_2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency.
Source: Author's calculations using data from Eguchi (2017).

***Figure 6*. Comparison of the external-shape dynamics of CEEI and DEEI: Nonparametric densities**



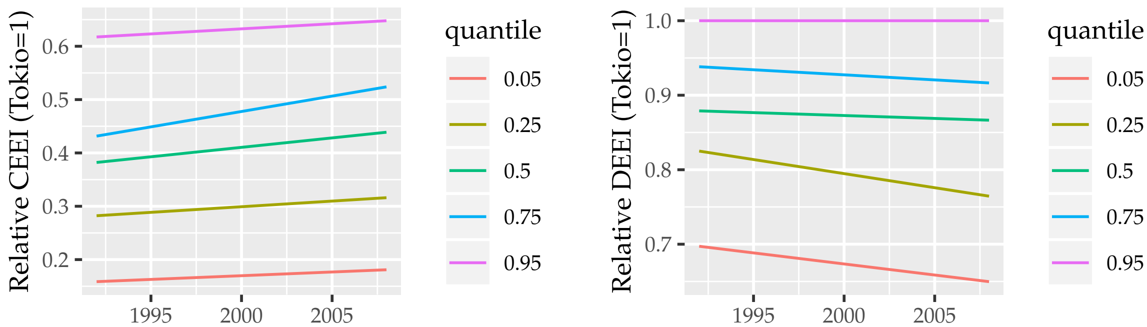
Note: CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency.
Source: Author's calculations using data from Eguchi (2017).

ever, increases overtime, particularly towards its left tail. Note that the standard sigma convergence analysis of Table 1b pointed out that the dispersion of the DEEI was increasing over time, and in sharp contrast with the CEEI, the results of the F-test for the DEEI were very close to be statistically significant at the conventional level of 10 percent.

4.2 Intra-Distributional Dynamics

Figure 7 provides a comparison of the intra-distribution dynamics of each environmental efficiency indicator through the lens of estimated quantile trends. In the case of CEEI, the analysis of the previous sections suggested a constant dispersion with a minor modal shift towards the upper tail. Panel (a) of Figure 6 not only confirms these findings, but also indicates that this modal shift is related with the dynamics of the 0.75 quantile. Also, relative to other quantiles, the slope of the 0.75 quantile is considerably steeper. Thus, only the regions from this quantile are showing some signs of convergence. In sharp contrast with this pattern, most of the quantiles of DEEI show negative slopes, which indicates intra-distributional divergence, which is particularly more evident in the two bottom quantiles.

Figure 7. Comparison of the intra-distribution dynamics of CEEI and DEEI: Quantile trends



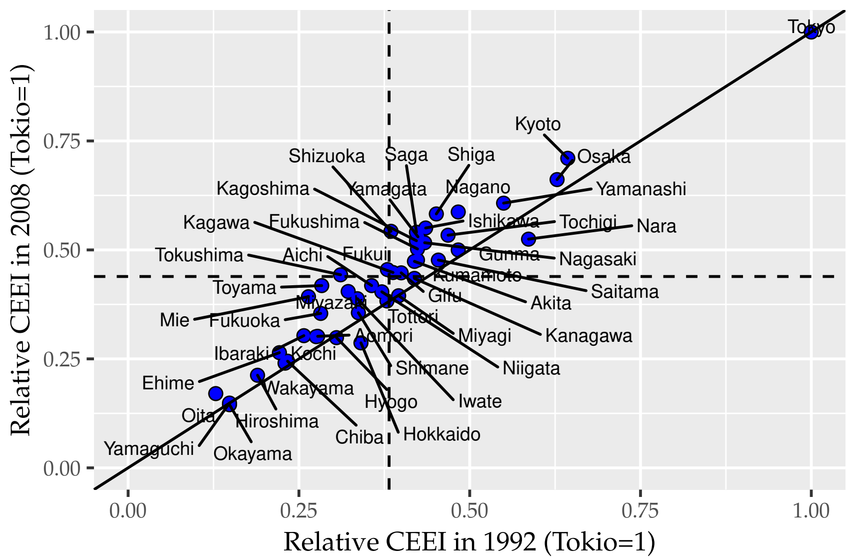
(a) (b)

Note: Since the data is only available for two years (1992 and 2008), a simple linear interpolation is applied to estimate the values of each intermediate year. CEEI stands for conventional environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator.
Source: Author's calculations using data from Eguchi (2017).

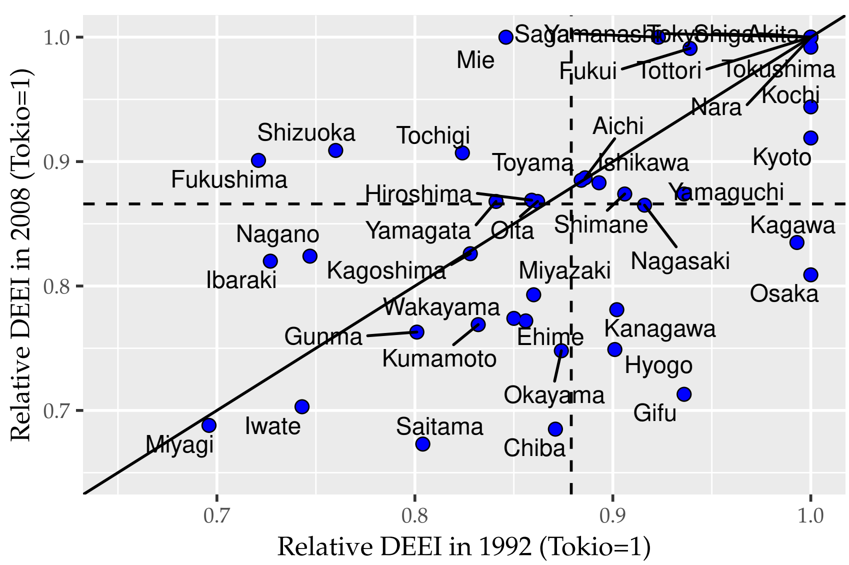
Figure 8 provides an alternative comparison of the intra-distribution dynamics of each indicator using mobility scatterplots. The advantage of this approach is that we can study the mobility of each region instead quantile groups. In the case of CEEI, the most noticeable feature is that all regions are lining up very close to a 45-degree line (Figure 8, Panel (a)). This pattern indicates relative stagnation in the worst case or very limited upward mobility (i.e., catching up with Tokio) in best case scenario. The dashed lines indicate the value of median observation in each year; they help us identify patterns of mobility in relation to the progress of the median region. Since almost all regions are located either in the lower-left or upper-right quadrant, one can conclude that the regions that were below (above) the median in the year 1992, remained below (above) it in the year 2008.

Panel (b) of Figure 8 shows a much more heterogeneous performance in the case of the DEA-based environmental efficiency indicator (DEEI). Regions are far away from the stagnation (45-degree) line. In fact, most of them are below it—including the median region. This downward mobility pattern reinforces previous findings regarding a tendency towards regional divergence that is more evident in the case of DEEI.

Figure 8. Comparison of the intra-distribution dynamics of CEEI and DEEI: Mobility Scatterplots



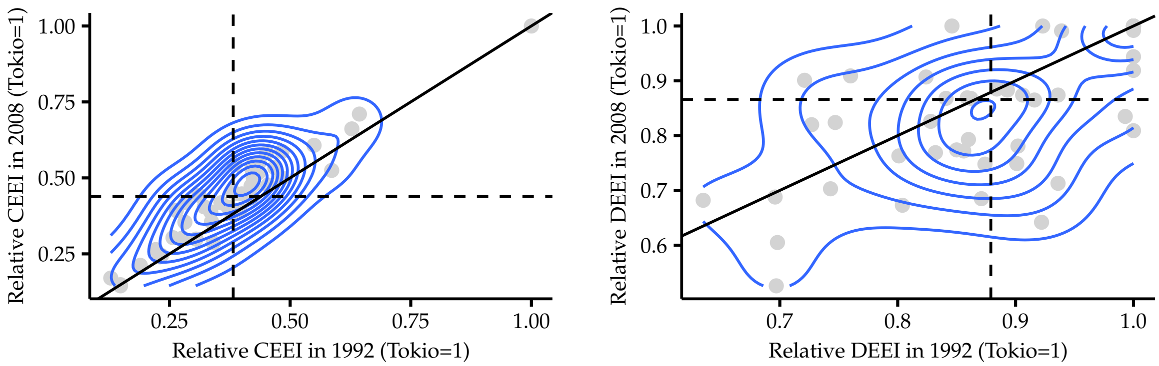
(a)



(b)

Note: CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency.
Source: Author's calculations using data from Eguchi (2017).

***Figure 9.* Comparison of the intra-distribution dynamics of CEEI and DEEI: Bivariate nonparametric distribution**



(a) (b)

Note: CEEI stands for "conventional" environmental efficiency indicator and DEEI stands for DEA-based environmental efficiency indicator. The word "conventional" refers to the ratio the ratio of gross regional production to CO2 emissions. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency.
Source: Author's calculations using data from Eguchi (2017).

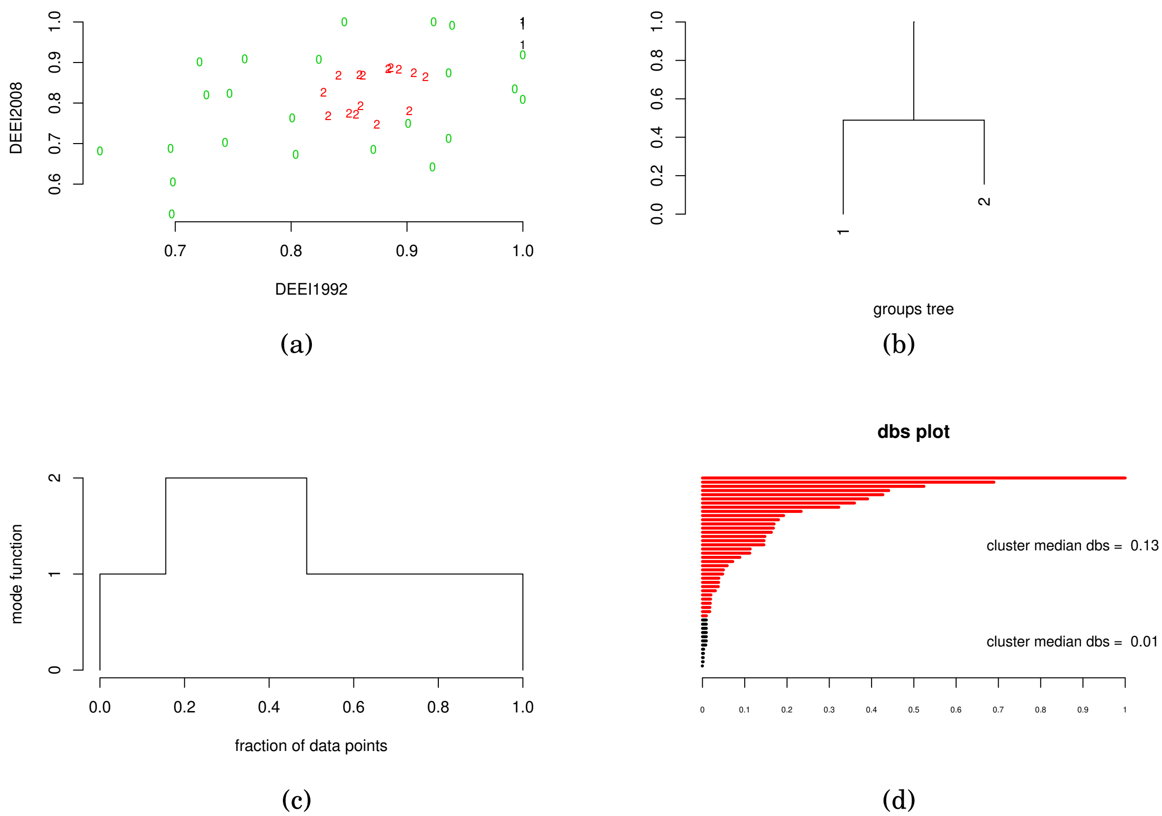
4.3 Regional Convergence Clusters

Figure 9 builds on the findings of the previous mobility scatterplots. Through the lens of a nonparametric bivariate distribution, it shows the areas of high density of each environmental efficiency indicator. In the case of CEEI, there is only one area of high density that is just above the median of each period. The DEEI, however, shows two distinct areas of high density along the stagnation (45-degree) line. The first one is located just under the median and the second one is located around the value of Tokio.

Figure 10 further evaluates the existence of two separate density areas for the DEA-based environmental efficiency indicator (DEEI). In the convergence literature, the existence of multiple areas of high density along the stagnation (45-degree) line is commonly interpreted as suggestive evidence of convergence clubs or clusters. Using the density-based clustering algorithm of Azzalini et. al (2007) and Menardi and Azzalini (2014), it is possible to identify two core clusters of convergence. Panel (a) of Figure 10 shows both the regions that belong to each cluster of convergence (market with the number 1 or 2) and the regions that are still in a transitional stage (market with the number 0).

In terms of the size and composition of each group, the superior cluster,

***Figure 10*. Density-based cluster analysis for the DEA-based environmental efficiency indicator (DEEI)**



Note: DEEI stands for DEA-based environmental efficiency indicator. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency.
Source: Author's calculations using data from Eguchi (2017).

which one may call it the "Tokio-led" cluster, is composed by 8 regions. On the other hand, the lower cluster, which one may call it the "median-led" cluster, is composed by 15 regions. Finally, the group of transitioning regions is composed by 23 prefectures.

The rest of the panels of Figure 10 provide additional clustering diagnostics that help us evaluate the possible destination of the transitioning regions. As suggested by the mode function (Panel c) and the Density Based Silhouette plot (Panel d), most of the transitioning regions would end up in the "median-lead" cluster. This cluster would increase its size from 15 regions to 34 regions. On the other hand, the "Tokio-led" cluster would only receive 4 additional members, ending up with 12 regions.

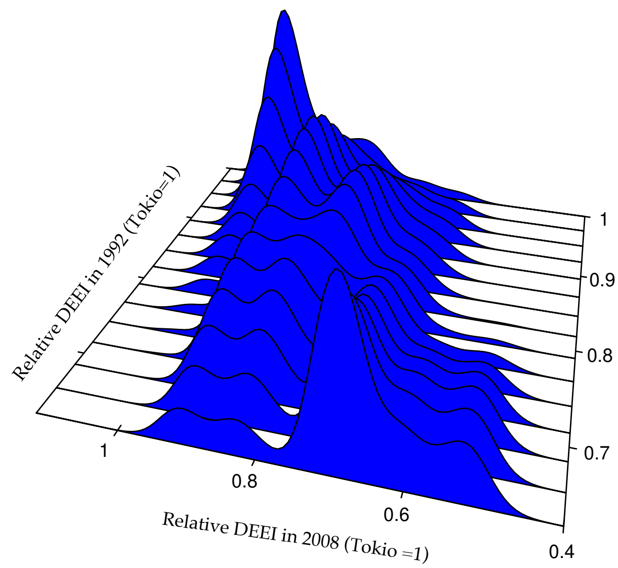
Finally, Figure 11 shows the conditional distribution version of the previously estimated intra-distribution dynamics. Through the lens of the conditional distribution estimator of Hyndman et. al(1996), we can observe the evolution of the environmental efficiency distribution, given its initial level in the year 1992. Also, in this modeling framework, both the "Tokio-led" and the "median-led" clusters are clearly identifiable. However, two additional "bottom" clusters are also identified (See Panel b).

It is important to clarify and reconcile the difference in the number of clusters between the two methodologies. The framework of Azzalini et. al (2007) and Menardi and Azzalini (2014) identifies modal clusters in two stages. In the first stage, the estimation of an un-conditional bivariate distribution suggest the existence of two modes (Figure 10b and 10c), which are then used to map two core clusters (Figure 10a and 10d). Transitioning regions that were not originally part of the core clusters are then *allocated only to those two clusters* in a second stage.[12] The framework of Hyndman et. al (1997), on the other hand, estimates conditional-bivariate distribution that suggest the existence of multiple modes *depending on the initial level of the variable* (Panel a).

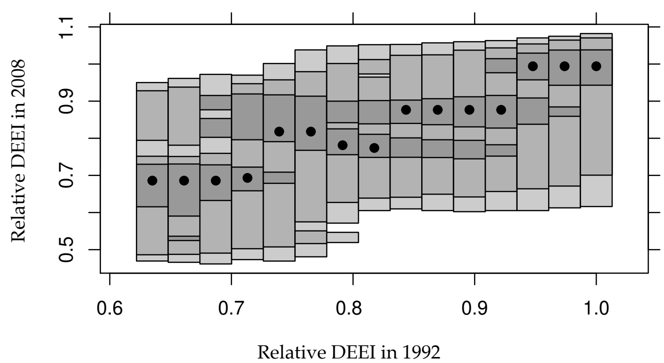
Although these two frameworks differ in the allocation mechanism of the transitioning regions, they share a common element: the unconditional bivariate kernel distribution. This distribution is the main criteria for finding the clusters of Menardi and Azzalini (2014). It is also a necessary input to estimate the conditional distribution of Hyndman (1996). Thus, to reconcile the results, the Azzalini approach could be suggesting a lower bound (where the initial conditions of the regions are not binding), whereas the Hyndman approach could be suggesting an upper bound (where the initial conditions of the regions are binding). More intuitively, *not all* transitioning regions may necessarily end up in any of the two superior convergence clusters. Specially some less developed regions may end up in a new equilibrium that is below the median. The approach of Hyndman emphasizes this possibility by pointing out the existence of two separate areas of high density around the lower tail of the conditional distribution.

[12]The statistical criteria for this allocation exploits recent advances in computational geometry, in particular the identification of a Delaunay triangulation and Voronoi tessellation.

Figure 11. Nonparametric conditional distribution for the DEA-based environmental efficiency indicator (DEEI)



(a)



(b)

Note: DEEI stands for DEA-based environmental efficiency indicator. DEA stands for Data Envelopment Analysis and it is a nonparametric analysis of relative efficiency. Panel (a) shows a stacked conditional density (SCD) plot that is conceptually similar to a Markovian transition matrix. For a given level of environmental efficiency in 1992, the SCD traces its evolution and illustrates its final configuration in the year 2008. Panel (b) shows high density region (HDR) plot. The darker-shaded stripes indicate higher density areas, and the mode for each conditional density is shown as bullet point.
Source: Author's calculations using data from Eguchi (2017).

5 Concluding Remarks

Both increasing the national (average) level of environmental efficiency and reducing its cross-regional dispersion are important means to reduce CO2 emissions. Although Japan has made considerable progress on the former, it has made much less progress on the latter. In this context, this paper studies environmental efficiency dispersion across 46 prefectures in Japan over the 1992-2008 period. In particular, through the lens of both classical and distributional convergence frameworks, this paper quantifies progress in regional convergence by using the two alternative measures of environmental efficiency suggested by Eguchi (2017). The first measure is a conventional environmental efficiency indicator (CEEI for short) that focuses only on the (desirable and undesirable) outputs of the production process. It is constructed as the ratio of gross regional product to CO2 emissions. The second measure (DEEI for short) is based on the data envelopment analysis model, and it includes not only desirable and undesirable outputs, but also inputs such as buildings, roads, private capital stock, and the number of employees.

The first empirical regularity documented in the paper is that the relation between CEEI and DEEI is highly non-linear. Interestingly, there appears to be a U-shaped relationship between these two indicators of environmental efficiency. This pattern could suggest that although initial improvements in environmental efficiency (higher CEEI) are associated with relatively inefficient use of production inputs (lower DEEI), regions could also learn how to use their production inputs more efficiently (higher DEEI) as they produce more output (higher CEEI).

The classical analysis of beta and sigma convergence suggests that—on average—environmental efficiency differences across regions have not decreased during the 1992-2008 period. Low-efficiency regions are not catching-up with the high-efficiency regions in an statistically significant way (i.e., lack of beta convergence). Moreover, across multiple indicators of dispersion, environmental differences across prefectures have actually increased (i.e., lack of sigma convergence).

The next section of the results goes beyond these two classical measures of convergence. Through the lens of a series of nonparametric distributional methods, it evaluates the evolution of the entire environmental efficiency distribution. This analysis points out three central aspects of distributional convergence: (1) the dynamics and composition of the external shape of the distribution; (2) the intra-distributional dynamics; and (3) the formation of local clusters of convergence within the distribution.

As expected, there are considerable distributional differences between the two indicators of environmental efficiency. On the one hand, the cross-regional distribution of CEEI shows a constant spread with a unique mode slightly shifting to the right. On the other hand, the distribution of DEE shows an increasing spread from its lower tail and a drastic transformation of its shape from a trimodal distribution into a left-skewed more unimodal distribution.

There are also sharp differences in terms of intra-distributional dynamics. In the case of the CEEI, all regions are lining up very close to the stagnation line of a mobility scatterplot. In contrast, in the case of the DEEI, most regions are far away from the stagnation line. Also, consistent with the the lack of sigma convergence finding, backward mobility appears to be the dominant force within this distribution.

Given the lack of overall convergence and the particularly large degree of backward mobility within the DEEI distribution, the last subsection of the paper evaluates the existence of local convergence clusters for the DEA-based indicator of environmental efficiency. The *unconditional* distribution framework of Azzalini et. al (2007) and Menardi and Azzalini (2014) suggests the existence of tree groups of regions: a “Tokio-led” convergence cluster (composed by 8 prefectures), a “median-led” convergence cluster (composed by 15 prefectures), and a group of transitioning regions (composed by 23 prefectures). Moreover, further results from the *conditional* distribution framework of Hyndman et. al (1996) suggest that it is possible that some of those transitioning regions may end up converging to two additional bottom clusters. Thus, the cross-regional dynamics of the DEA-based environmental efficiency indicator (DEEI) may be characterized by at least two convergence clusters.

To sum up, all the previous results highlight the lack of overall regional convergence in both indicators of environmental efficiency. However, important differences exist between these indicators. Perhaps the most important is that the DEA-based environmental efficiency indicator (DEEI) help us identify local convergence clusters. This identification is potentially driven by two sources. The DEA framework itself, which is used to aggregate the inputs and related them to outputs. And the inclusion of production inputs in the calculation of the environmental efficiency index. If local convergence clusters are mostly driven by the latter, then there is clear policy implication: environmental policy efforts—focalized at the cluster level—should aim at improving the efficiency with which production inputs are used.

Finally, further research on environmental efficiency convergence in Japan seems promising in at least two immediate fronts. First, one could exploit some of the recent advances in the spatial analysis of convergence and evaluate to what extend the spatial neighbors of a particular region help accelerate or retard its rate of convergence.[13] Second, one could also exploit some recent advances in dynamic panel data analysis. In particular, the convergence test and clustering framework of Phillips and Sul (2007; 2009) could be useful to evaluate how the cluster composition behaves over longer time horizons.

References

Arbia, G., Basile, R., and Piras, G. (2006). Analyzing Intra-Distribution Dynamics: A Reappraisal. ERSA conference papers ersa06p262, European Regional Science Association.

Barro, R. J. and Sala-I-Martin, X. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, 1991(1):107–182.

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

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

Baumol, W. J. (1986). Productivity growth, convergence, and welfare: what the long-run data show. *The American Economic Review*, pages 1072–1085.

Bianco, S. D. (2016). Going clubbing in the eighties: convergence in manufacturing sectors at a glance. *Empirical Economics*, 50(2):623–659.

Eguchi, S. (2017). Accounting for resource accumulation in japanese prefec-

[13]See, for instance, the survey of article of Rey and Le Gallo (2009) or the work of Maza and Hierro (2012).

tures: An environmental efficiency analysis. *Journal of Economic Structures*, 6(1):16.

Epstein, P., Howlett, P., and Schulze, M.-S. S. . M. (2003). Distribution dynamics: stratification, polarization, and convergence among oecd economies, 1870–1992. *Explorations in Economic History*, 40(1):78–97.

Henderson, D. J. and Parmeter, C. F. (2015). *Applied Nonparametric Econometrics*. Cambridge University Press, Cambridge.

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

Magrini, S. (2004). Regional (di) convergence. In *Handbook of regional and urban economics*, volume 4, pages 2741–2796. Elsevier.

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

Maza, A., Hierro, M., and Villaverde, J. (2012). Income distribution dynamics across european regions: Re-examining the role of space. *Economic Modelling*, 29(6):2632–2640.

Menardi, G. and Azzalini, A. (2014). An advancement in clustering via nonparametric density estimation. *Statistics and Computing*, 24(5):753–767.

Mendez-Guerra, C. (2017). Heterogeneous growth and regional (di)convergence in Bolivia: A distribution dynamics approach. *Conjunctural Economics*, 2(4):81–108.

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

Ministry of the Environment, J. (2013). The 3rd fundamental plan for establishing a sound material-cycle society.

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

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

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, pages 427–443.

Quah, D. T. (1996). Twin peaks: growth and convergence in models of distribution dynamics. *Economic Journal*, pages 1045–1055.

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

Rey, S. J. and Le Gallo, J. (2009). Spatial analysis of economic convergence. In *Palgrave handbook of econometrics*, pages 1251–1290. Springer.

Sala-i Martin, X. (1996). The classical approach to convergence analysis. *Economic journal*, pages 1019–1036.

Tanikawa, H., Fishman, T., Okuoka, K., and Sugimoto, K. (2015). The weight of society over time and space: A comprehensive account of the construction material stock of japan, 1945–2010. *Journal of Industrial Ecology*, 19(5):778–791.

Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing*. Chapman and Hall, London.