Working Papers

Institute of Mathematical Economics

455

November 2011

Dealing with heterogeneity, nonlinearity and club misclassification in growth convergence: A nonparametric two-step approach

Harry Haupt and Verena Meier

IMW · Bielefeld University
Postfach 100131
33501 Bielefeld · Germany

email: imw@wiwi.uni-bielefeld.de
http://www.imw.uni-bielefeld.de/research/wp455.php
ISSN: 0931-6558

Dealing with heterogeneity, nonlinearity and club misclassification in growth convergence: A nonparametric two-step approach.[*][†]

Harry Haupt and Verena Meier

Bielefeld University, Germany

Summary. Classical growth convergence regressions fail to account for various sources of heterogeneity and nonlinearity. While recent contributions are able to address either the one or the other, we present a simple two-step method to address both issues. Based on a slightly augmented version of a recently proposed algorithm to identify convergence clubs, we formulate a flexible nonlinear framework which allows to analyze convergence effects on both individual and club level, while alleviating potential misclassification in the club formation process using simultaneous smoothing over the club structure. The merits of the method are illustrated for data on different aggregational levels.

[*]We are indebted to Joachim Schnurbus for many helpful comments. Of course all remaining errors are our own responsibility.

[†]Address correspondence to: Harry Haupt, Centre for Statistics, Bielefeld University, PO Box 100131, 33501 Bielefeld, Germany. E-mail: hhaupt@wiwi.uni-bielefeld.de.

1 Convergence, heterogeneity, and nonlinearity

Classical econometric convergence analysis in the sense of Mankiw et al. (1992), Barro et al. (1991) and Barro and Sala-i-Martin (1992) is based on the concept of absolute β-convergence, the latter meaning that poor countries or regions grow faster than rich ones. The workhorse is a linear regression model for cross-sectional data

$$\log(y_{i,t}) - \log(y_{i,0}) \stackrel{\text{def}}{=} g_{i,t} = \alpha - \beta \log(y_{i,0}) + u_{i,t}. \tag{1}$$

For each region i, $1 \leq i \leq N$, the growth rate of per capita income $g_{i,t}$ in region i – calculated over an a priori fixed time period 0 to t – is explained by its respective initial log per capita income $\log(y_{i,0})$. In this framework $u_{i,t}$ is assumed to be an idiosyncratic error term. Following from economic theory (see Barro and Sala-i-Martin, 2004), the parameter of paramount interest β is defined as $\beta \stackrel{\text{def}}{=} 1 - \exp(-\rho t)$, with convergence parameter ρ measuring the speed of convergence. Absolute β-convergence is assumed if $\rho > 0$ and consequently $\beta > 0$. Then countries with a smaller per capita income grow faster than rich countries such that income differences decrease[1].

Two strands of criticism confront classical growth convergence analysis. First of all, the concept of β-convergence could be invalid due to neglected heterogeneity, that is the relationship between $g_{i,t}$ and $\log(y_{i,0})$ in (1) may depend on parameters indexed by cross-section (for fixed t), say $\alpha_{i,t}$ and $\beta_{i,t}$ (e.g., Masanjala and Papageorgiou 2004, Alfò,Trovata and Waldmann 2008, Canarella and Pollard 2004). Second, the lack of functional flexibility of the classical parametric formulation and estimation of (1) and hence the potential of neglected nonlinearities (e.g., Kalaitzidakis et al. 2001, Liu and Stengos 1999, Maasoumi, Li and Racine 2007, Quah 1993,1997, Henderson 2010).

[1] Please note that all of the considerations in this paper can be readily applied to models of conditional convergence, e.g., Haupt and Petring (2011).

Addressing the first issue, Durlauf and Johnson (1995) and Canova (2004) use (1) within groups of countries with common convergence behavior in the group and heterogeneous behavior between groups. Phillips and Sul (2003, 2007a,b, 2009), hereafter PS, argue that classical convergence analysis based on (1) is prone to deliver inconsistent results and invalid convergence tests due to potential heterogeneity in the convergence parameter β over time, countries, and individual technology levels. PS show that due to omitted heterogeneity the error term in (1) includes endogenous variables and variables which are correlated with dependent and independent variables. As a remedy PS suggest to enable a variation of the transition parameter and growth rate over districts and time[2]. They propose a nonlinear dynamic factor model

$$\log(y_{i,t}) = a_{i,t} + x_{i,t}t = \left(\frac{a_{i,t} + x_{i,t}t}{\mu_t}\right)\mu_t \stackrel{\text{def}}{=} b_{i,t}\mu_t, \qquad (2)$$

where $x_{i,t}$ is an individual technology process parameter, $b_{i,t}$ is the idiosyncratic time-varying element and μ_t a common trend factor measuring global technological progress.

Then $b_{i,t}$ can be interpreted as the transition path of economy i to the global growth path μ_t and is calculated as the log per capita income of district i in period t. By eliminating the global growth component, the relative transition path

$$h_{i,t} = \log(y_{i,t})/N^{-1}\sum_{i=1}^{N}\log(y_{i,t}) = b_{i,t}/N^{-1}\sum_{i=1}^{N} b_{i,t}$$

measures the transition element for economy i in period t in relation to a cross-section average. Then global convergence — all countries have the same fraction

[2]Note that heterogeneity of parameters in (1) may also occur across the conditional distribution of the growth rates g_{it}. Haupt and Petring (2011) apply quantile regression estimation and test but do not find empirical evidence in favor of such types of heterogeneity using the data from Mankiw et al. (1992). Hence this issue will not be pursued here.

of global per capita income — is assumed to be present if

$$h_{i,t} \to 1, \quad \text{for all } i, \quad \text{as } t \to \infty. \tag{3}$$

The log t regression of Phillips and Sul (2007a,b, 2009)

$$\log\left(H_0/H_t\right) - 2\log(\log(t)) = a + \gamma \log(t) + u_t \tag{4}$$

now tests (3) using the mean square transition differential $H_t = N^{-1}\sum_{i=1}^{N}(h_{i,t} - 1)^2$. In case of global convergence $H_t \to 0$ as $t \to \infty$. The authors show that $H_t \sim A/\log(t)^2 t^{2\alpha}$ as $t \to \infty$, where $A \geq 0$ is a constant and α equals the rate of cross-section transition variation dissolving over time. Under the null hypothesis the regressor diverges to ∞ and under the alternative the regressor diverges to $-\infty$. A negative value, however, does not necessarily imply that there is divergence but that there may exist some convergence clubs instead of global convergence. Using a one-sided t-test we test the null hypothesis of $\gamma \geq 0$.

For comparing the concept of β-convergence and the concept of PS we have to concentrate on the initial and final period because β-convergence doesn't consider intermediate periods. Independently from their initial income in 0, all countries or regions have the steady-state per capita income in T (the mean income of T) if convergence in this sense is fulfilled over the period from $[0, T]$. In empirical samples we assume that the points don't lie exactly on a line but they spread sparsely around the mean. Specifying the relation in a linear regression model yields to

$$\log(y_{i,T} = \alpha + b\log(y_{i,0}) + u_{i,T}, \tag{5}$$

where $b = 0$ under convergence in per capita hypothesis. Assuming $b = 1 - \beta$ (5) is equal to classical β-convergence modeling. Thus, convergence in the sense of PS is a special case of β-convergence where $\beta = 1$.

The problem of potential nonlinearities in growth (and convergence) regressions has been addressed recently by applying fully nonparametric methods for regression and specification testing (e.g., Haupt and Petring, 2011, and the literature cited therein). The authors use a local linear kernel estimator with a generalized product kernel function proposed by Racine and Li (2004) and Li and Racine (2008). Using the data from Mankiw et al. (1992) they find considerable evidence for parametric misspecification and a superior performance of a nonparametric model, though no evidence for heterogeneity across the growth distribution. Using the data from Phillips and Sul (2009) in this paper we show that nonlinearities still may pose a problem, even on club level.

The contribution of this paper is to propose a new method to simultaneously address nonlinearity and heterogeneity in a convergence club model. In order to allow for heterogeneity, model (1) is extended to allow for club-specific convergence parameters. The resulting model is tested for parametric misspecification against a fully nonparametric alternative. Whenever the former linear model is rejected, the latter nonlinear model is used to estimate convergence effects. Basically, to pursue such an approach we apply two steps: First we discuss potential pitfalls, slightly augment the clubbing algorithm of PS and address issues arising from potential misspecification of their approach for testing the convergence condition (3). Second, we formulate a flexible nonlinear regression framework for studying growth convergence allowing for the estimation of club specific convergence effects which — in contrast to models on the club-level — do not suffer from potential errors in the classification of members and non-members.

The remainder of the paper carefully describes the proposed two-step procedure and its application to various data sets. In Section 2 we review the clubbing algorithm of PS, discuss potential pitfalls, and suggest a simple remedy for a potential problem of the algorithm (step 1). Section 3 discusses the specification search of a club-level convergence regression model and estimation of fully nonparametric regressions (step 2). The potential problem of club misclassification and consequences for regression inference are highlighted in Section 4. Finally, in Section 5 we apply the proposed method to study data on different levels of aggregation from the countries of the Penn World Tables, Japanese prefectures, and districts from reunified Germany, respectively.

2 Convergence club identification

To identify convergence clubs PS use a clubbing algorithm consisting of four steps

<1.> (Cross-section ordering): Order countries according to the $\log(y_{i,t})$ in final period.

<2.> (Form a core group of k^*, $2 \leq k^* < N$, countries):

- <2.1> Find the first two highest successive countries for which the log t test statistic $t_k \geq -1.65$. If the condition does not hold for any $k = 2$, drop the country with highest $\log(y_{i,t})$ and restart the procedure with the remaining countries.

- <2.2> Start with the $k = 2$ countries identified in 2.1, increase k proceeding with the subsequent country

  from order, run the log t regression, and calculate t_k. Stop increasing k if convergence hypothesis fails to hold (i.e. $t_k < -1.65$). Take the k^* countries with the highest test statistic from all k countries satisfying the convergence hypothesis for core group.

<3.> (Sieve the data for new club members):

- <3.1> Form a complementary core group with all remaining countries.

- <3.2> Add one country at a time from the complementary core group to the core group, run the log t regression, add the country to a club candidate group if the convergence test statistic is greater than a critical value $c^* = 0$. Form a convergence club of the candidate group and the core group.

<4.> (Recursion and stopping rule): Form a second group from all countries which fail the sieve condition in step 3 and run log t regression. If the convergence hypothesis cannot be rejected, all remaining countries form a new convergence club. Otherwise, for the remaining countries start again with step 2 for finding a new k^*.

<5.> (Club merging): Run $\log t$ regression for all groups of subsequent clubs. Merge those clubs fulfilling the convergence hypothesis commonly.

Composing the clubs in accordance to this algorithm does not ensure that the convergence hypothesis holds for each respective club. PS (2007) are aware of this problem and propose to increase the critical value c^* for raising the power of the corresponding test. Such a remedy, however, does not work in general, for instance for the German district-level data discussed in Section 5.2 or when we replace the initial cross-section ordering rule by an (equally plausible) alternative, Thus, we may want to augment step <3.> of the algorithm in a way such that convergence is assured using a data-based criterion.

<3.3> If the countries from core and candidate group hold convergence hypothesis commonly, go to step 4. If not, form a convergence club with the candidate country with highest test statistic and the core group. Add one candidate country at a time to convergence club, run $\log t$ regression and add the country with highest test statistic to the convergence club. Continue adding new countries to the convergence club until no further candidate country fulfills convergence hypothesis.

In all empirical applications discussed in Section 5 we find that the shape of observed points in the $\log t$ regression (4) to be parabolic and convex for convergence clubs. This is due to the construction of the regressor. Under the null hypothesis H_t converges to zero as $t \to \infty$ as a monotonically decreasing convex function. Calculating H_0/H_t inverts this shape into a monotonically increasing convex function. Taking the logarithm damps the curvature or even linearizes the curve. The second part of the regressor $2\log(\log(t))$ is a monotonically increasing concave curve. Subtracting this second concave part from the first convex/linear curve leads to a parabolic and convex trajectory. Thus, under the null we expect a nonlinear regression relationship. Those results suggest that the interpretation of the $\log t$ regression should be handled with care.

3 Estimation of club-based convergence regression

We want to apply a classical convergence analysis in the sense of Mankiw et al. (1992) while allowing for data-driven heterogeneity and nonlinearity. Thus, in a first step, we assign the regions to clubs using the algorithm discussed above. In a second step we include a categorical club variable $club_i$ in (1)via the j dummy variables $club_{i,j}$ which are equal to 1 if country i is in club j. The resulting baseline model allows to estimate a regression line $\delta_j + \pi_j \log(y_{i,0})$ for every club j, $1 \leq j \leq m$, i.e.

$$g_{i,t} = \sum_{j=1}^{m} \delta_j club_{i,j} + \sum_{j=1}^{m} \pi_j \log(y_{i,0}) \cdot club_{i,j} + u_{i,t}. \tag{6}$$

In contrast to the classical convergence model (1), the baseline model (6) allows for a considerable degree of heterogeneity. However, there are very small clubs for several applications and thus the interpretation of the parameters for those clubs should be handled with care. The main point of criticism, however, is that this model may suffer from potential misclassification of the club composition (see Section 4). Furthermore, the model does not allow for further nonlinearities.

In order to address the problem of potential nonlinearities we can employ a fully nonparametric alternative

$$g_{i,t} = G(\log(y_{i,0}), club_i) + u_{i,t}. \tag{7}$$

This approach allows to estimate not only club-level effects — which Durlauf and Johnson (1995) interpret to represent averages of the underlying individual effects for each country — but further nonlinearities. The approach allows to consider mixed data with both continuous (here: initial income $\log(y_{i,0})$) and categorical (here: $club_i$, an ordered categorical variable) covariates.

In model (7) for every $\log(y_{i,0})$ a linear model for its direct neighborhood of size λ is estimated. The latter regression can be estimated by a local linear kernel estimation with general regression function $G()$. The points get different weights given by a weighting function $W()$, the generalized product kernel presented by Racine and Li (2004)

$$W(z_0, z_i, \lambda) = \prod_{k=1}^{2} w_k(z_{0k}, z_{ik}, \lambda_k)$$

The idea is that all types of covariates are smoothed with a certain weighting (kernel) function corresponding to the scale level of the covariate, assigning an individual smoothing parameter (the bandwidth) to each covariate. For the continuous variable $\log(y_{i,0})$ a second order Gaussian kernel $w_k(z_{0k}, z_{ik}, \lambda_k) = {\lambda_k}^{-1}\phi\left(z_{ik} - z_{0k}/\lambda_k\right)$ is used, where ϕ is the standard normal density and $\lambda \in (0, \infty)$. Due to the Gaussian distribution, points close to $\log(y_{i,0})$ ($z_{ik} \approx z_{0k}$) get higher weights than points lying near the boarder of the neighborhood. For the ordered categorical variable $club_i$ we use a kernel function of Racine and Li (2004) $w_k(z_{0k}, z_{ik}, \lambda_k) = \lambda_k^{|z_{ik}-z_{0k}|}$. The bandwidth λ_k lies in the interval $[0, 1]$. For a value of $\lambda_k = 0$, the kernel $w_k(z_{0k}, z_{ik}, 0)$ is an indicator function for category z_{ik} and for $\lambda_k = 1$ the kernel function is constant over all categories. Thus, the variable is irrelevant. In contrast to a classical (frequency) approach, the obvious advantage of this method is potential smoothing of categorical variables. Hence, we are even able to estimate convergence behavior for small club sizes.

The resulting minimization problem is a weighted local least squares problem

$$\min_{\alpha(z_0),\beta(z_0)} \sum_{i=1}^{n} [g_{i,t} - \alpha(z_0) - (\log(y_{i,0}) - \log(y_{0,0}))\beta(z_0)]^2 W(z_0, z_i, \lambda),$$

where $z = (\log(y_{i,0}), club_i)$ is the vector of covariates and the "local" part is considered by the dependence of the parameters α and β on z_0. The remaining problem consists of finding optimal values for the bandwidth vector λ. We obtain λ using a data-driven least squares cross validation approach, where we minimize the objective function $CV(h) = n^{-1}\sum_{i=1}^{n}(g_{i,t} - \hat{G}_{-i}(z_i))^2 M(z_i)$, and $\hat{G}_{-i}(z_i)$ is the leave-one-out kernel estimator of regression function, and M is a weighting function bounded between 0 and 1, usually set to $M = 1$ (see Li and Racine, 2004).

4 Assessing potential club member misclassification

4.1 Sensitivity of results with respect to the clubbing algorithm

In contrast to model based clubbing algorithms (e.g. Juárez and Steel, 2010), the method of Phillips and Sul (2007a,b, 2009) discussed in Section 2 does not provide estimates of the misclassification probabilities for each club member. A first step towards exploring potential classification error is to check for hints on the existence of positive error probabilities by inspecting whether the "selection of core groups is robust to initial data orderings" (see Phillips and Sul, 2009, footnote 11, p. 1170). Considering the problem of an unknown true ordering rule (see Canova, 2004) we try different concepts in step <1.> and check whether considerable differences in club composition are obtained. This indicates large uncertainties which should be addressed in empirical convergence analysis.

As alternatives to the amount of final period income (final ordering), hereafter denoted as ordering rule (I), as used by Phillips and Sul (2009) we employ the following. (II) Order corresponding to the average income of all years (average ordering) for capturing potential time series volatility. Phillips and Sul (2007) propose to average over the last fraction of the sample to ensure a higher influence of recent periods. (III) Another alternative is ordering according to the difference between final period income and income in first period, capturing the income change over time (difference ordering). (IV) Finally, combining the ideas on the final period and capturing volatility, a decreasing weights ordering is employed.

We note that in all of our applications discussed below the use of different ordering rules leads to considerable differences in club sizes and composition, respectively. For evaluating the empirical performance we compare the out-of-sample performance of the convergence regression models for ordering rules (I)-(IV). As an a posteriori selection criteria for ordering rules we run a cross-validation (e.g., Haupt and Petring, 2011) and choose the model with the smallest average squared error of prediction.

4.2 Addressing misclassification in nonparametric regression

While there is no obvious remedy for the misclassification problem in the parametric model (6), the nonparametric model (7) may offer one. Data-driven bandwidth selection for the club variable deals with the question of uncertainty of club composition. Using the kernels proposed by Racine and Li (2004), the optimal estimated bandwidth is bounded between 0 and 1. A bandwidth of approximately 0 means that the influence of this variable is such that for estimating the function (7) for a club only observations from this club are used. This occurs when the functional form is sufficiently different with respect to the different clubs or if the observations show sufficiently different convergence behavior. We can interpret this in the sense that there is a rather low probability of misspecification, thus the clubs are well chosen. With increasing values of the bandwidth the error probability for club membership rises. If the bandwidth is considerably greater than 0, observations from all clubs are used to estimated regression functions for each club and thus, there is no influence of the variable. This suggests that there is evidence in favor of an only weak or even non-existent club structure. Thus, the bandwidth of the categorical club variable serves for an a posteriori quantification of the classification (and underlying error probabilities) as a whole.

By using the nonparametric approach including the club variable we obtain individual influences of each observation while considering the uncertainty with respect to club membership, instead of a single fixed convergence regression line for each club in the parametric approach. The club-structure on the other hand has the advantage of being backed up by economic theory. Although it may produce a faulty number and/or composition of clubs, the simultaneous smoothing of the continuous and the categorical variable is capable of alleviating this problem. In summary we include heterogeneity in the sense of Phillips and Sul, reduce uncertainty of club composition, and capture potential nonlinearities, and hence are able to address the main points of criticism of convergence regressions in recent literature.

4.3 Addressing misclassification in parametric regression

Given a set of data the initial problem a researcher faces is choosing either a parametric models such as (6) or a nonparametric model such as (7). In the context of mixed continuous and categorical covariates as in the present example this problem can be addressed by applying the test of Hsiao et al. (2007) (hereafter HLR test), which is based on the generalized product kernel estimator proposed by Racine and Li (2004) discussed in Section 3 above.

Using the same nonparametric configurations used for the nonparametric regression the HLR test checks if the parametric null model (6) is correctly specified. Whenever the HLR test rejects the null we apply the fully nonparametric model, enjoying the benefits discussed in the previous sections.

As the HLR test employs the bandwidths of the nonparametric regression, we are able to assess the error probabilities already after applying the test. Thus, if the test does not reject the parametric null hypothesis, we inspect the bandwidth λ_k of the cluster variable: If λ_k is close to zero, the parametric and nonparametric model work analogously and we may use the parametric model because there are no hints for club misclassification. If the bandwidth λ_k is greater than zero positive classification errors have a higher probability. In this case, however, we can still estimate a nonparametric model for the theoretical price of efficiency loss compared to the parametric model.

5 Applications featuring different levels of aggregation

In the following subsections the proposed method is illustrated with applications to three data sets based on different levels of aggregation — the countries from the Penn World Tables, the prefectures of Japan, and the districts from reunified Germany. These applications allow replication of our method and previous results in a wide sense. We use different levels of aggregation because we expect different levels of heterogeneity. Regions on district level come with similar technology and thus regions on this level converge to a similar or even the same steady state. This is the reason why the concept of absolute convergence is generally used for disaggregated data. However, different countries behave much more heterogeneously, because there are highly differing levels of technology. This is the reason why countries typically converge to different steady states. Classical convergence analysis captures this problem by extending (1) with additional covariates (e.g. investment rate, human capital) determining different steady states (see Sala-i-Martin, 1996). In our approach we use the concept of absolute convergence for all levels of aggregation because we allow for different steady states on club level by capturing the club variable. Independently from the aggregation level members of one club are assumed to offer homogenous convergence behavior and thus, we can assume similar steady states in a club. With respect to the aggregation level our empirical results reveal considerable differences in nonlinearity and heterogeneity, while we do not find clear evidence on the sensitivity of results with respect to the ordering rules discussed above[3].

[3]All computations in this paper are done using the software `R`, version 2.11.0, and version 0.40-4 of the `np`-package of Hayfield and Racine (2008). Of course, data and code are available from the authors.

5.1 Penn World Table country-level data

Using our two-step procedure we analyze convergence for Penn World Table (PWT) data of 152 countries over the years from 1970 to 2003. As global convergence is clearly rejected (p-value $\approx$ 0), the clubbing algorithm is applied. Table 1 displays parameter estimates and standard deviations before and after club merging for ordering rules (I)-(IV).

Final ordering (I) offers seven convergence clubs and no diverging countries[4] while one third of the countries are members of the first club. After merging six clubs remain. Using the other ordering rules we get different results. Average ordering (II) produces basically nine convergence clubs, but using club merging the number of clubs can be reduced to six clubs and the divergence group and similarly to final ordering, the first club is the biggest one and consists of 67 countries, while the other clubs are much smaller. The divergence group has six members. Difference ordering (III) produces only five non-mergeable convergence clubs (and one diverging country), while also 67 of the countries are members of the first convergence club. Decreasing weights ordering (IV) generates seven clubs which persist after merging. About half of the countries belong to the first club. In summary, the composition and number of convergence clubs seems to be highly sensitive with respect to the choice of the ordering rule.

The convergence behavior of the six clubs using final ordering (I) is displayed in Figure 1, where the relative transition coefficients are plotted against time. A closer look at the respective club members listed in Table 18) may raise some suspicion. For example Club 1 contains the USA and Botswana (e.g., Phillips and Sul, 2009). In absolute values the per capita income of the USA in 1970 is about 17429 US Dollars, compared to 1184 US Dollars in Botswana. Though in absolute values this gap rises considerably until 2003 (see Figure 2), in relative numbers it

[4] Using the same data, Phillips and Sul (2009) only identify five convergence clubs. For those clubs, however, we find the same parameter estimates and standard deviations.

decreases over time. While in 1970 the per capita income in Botswana is about 7% of the per capita income in the USA, in 2003 it is about 23%. Botswana also catches up in international comparison with respect to the relative transition coefficients h_{it}. In 1970 the per capita income in Botswana lies at about 80% (USA: 110%) of the cross-country average, while it rises to 91% in 2003 (USA: 105%). Thus, although the absolute incomes between these two countries differ extremely, the countries converge in the sense of Phillips and Sul as the respective h_{it} converge to 1.

In Figure 3 the box-plots of income in final period are displayed for the six clubs found by final ordering (I). While the incomes inside the clubs are close to each other, the income distribution between the clusters is very heterogeneous. For the same clubs in Figure 4 we display scatter plots of the log t regressions (4). The shape suggested by the trajectories in clubs 1 to 4 is parabolic and convex and thus may be interpreted in a way that in initial periods there are hints for divergence, while over the years we observe convergence because of a positive slope. For club 5 we detect more complex nonlinearities and convergence is assumed because γ is not significantly negative but the regressand decreases at the end of the period, indicating that there is no convergence[5]. Thus, to avoid a misinterpretation of the estimation and test results, the inspection of the log t regression scatter plots seems to be highly recommended.

The clustering algorithm may also be sensitive with respect to the respective time horizon. Thus, for the Penn World Table data we exemplarily analyze for final ordering whether number, size, and composition of clusters is constant for different time periods. We compare the results for the whole time horizon from 1970 to 2003 with consecutively shorter partial time spans, one from 1978 to 2003 and the other from 1986 to 2003. The reason why we choose both periods such that they also end in 2003 is that the income in the final period is the ordering

[5]Note that for other ordering rules and data sets, more clubs exhibit such a behavior.

criteria. Using the same final period enables to analyze how the length of the time horizon affects the cluster composition and the number of clusters.

Table 2 displays numbers, sizes, and compositions of clusters for the complete time horizon 1970 to 2003 and the period from 1978-2003, respectively. The cluster sizes for the complete time horizon is given in the last column containing the row sums, the clusters of the partial period 1978 to 2003 are given in the last row containing the column sums. For the partial horizon we find an additional convergence club and a divergence group. Although the number of clusters changed, their composition is quite stable as countries belonging to the first clubs over the complete horizon predominantly also are members of the first clubs in the partial horizon (and vice versa). As can bee seen from Table 3, the number of clusters rises to eight and one divergence group when comparing the shorter partial period from 1986 to 2003 to the complete time horizon. Again, though the number of clusters varies over time, the club composition seems to be quite stable over time. This results support the assumption that the club structure can be included as an ordered categorical variable when analyzing β-convergence.

Step one reveals hints for non-robust club sizes and club compositions with respect to ordering rules and time horzion as well as neglected nonlinearities in $\log t$ regressions. Both findings raise the question if potential misclassification of convergence clubs will affect estimation and inference within this framework. Thus, in the second step, we analyze the robustness of the club compositions resulting from the differing ordering rules (I)-(IV). For each ordering rule we estimate the parametric model (6) and the nonparametric model (7) and apply the HLR test. Finally, by running an out-of-sample cross validation analysis we select an optimal ordering rule according to lowest average squared error of prediction (ASEP).

The output for a classical β-convergence regression (1) is given in Table 4. The estimated convergence coefficient is negative, but there is no statistical significance. Advancing to the baseline model (6) including the club variable[6] suggests strong evidence for the existence of heterogeneity. The estimated coefficients are displayed in Table 5 and the resulting club-level regression lines can be seen in Figure 5[7]. The convergence coefficients are significant for convergence club one to five, but not for the sixth club (which consists of only two countries). The p-values of HLR tests (see column 3 of Table 6) are approximately equal to zero in all four cases suggesting the application of the nonparametric model (7). Table 6 displays the resulting bandwidth for nonparametric models. For ordering rules (I)-(IV) the estimated bandwidth for continuous regressor is smaller or equal to its standard deviation (1.09), respectively, also indicating a nonlinear influence of the regressor $\log(y_{i,0})$.

Overall, the clubs seem to be well chosen because the bandwidth of the club variable is very small independently from the respective choice of ordering rule. An out-of-sample cross-validation, however, offers a clear ranking for ordering rules with respect to ASEP. The pairwise comparison of the models given in Table 7 reveals that ordering rule (I) suggested by Phillips and Sul (2009) dominates all other ordering rules for the present data.

5.2 District-level data from reunified Germany

For this application data on the 439 German administrative districts are taken from the regional data base of the statistical agencies of Germany for per capita income measured as the GDP divided by the corresponding number of citizens for the years 1996 to 2005. The log t regression for German regional data suggests clear evidence against global convergence (p-value ≈ 0), but we are able to find

[6] We only present results for ordering rule (I) because later on we find that this rule performs best. The results for other ordering rules are similar.

[7] The estimated coefficients for quartile regression are similar to mean regression. Thus, there is no more heterogeneity over the conditional distribution of the regressand.

the clubs summarized in Table 8 for the four ordering rules and classification before merging and after merging. Again, difference ordering produces fewest number of clubs, only eight before and five after merging. Average and decreasing weights ordering reveal highest number of clubs, 24 before and ten respectively eight after merging. Notably, the first two clubs and the divergence group for almost all orderings except difference ordering are very small while for difference ordering the first club includes about 25% of data and the divergence group even one third. Figure 6 plots the relative transition coefficients over time for the convergence clubs and the diverging group of final ordering. The plots support the convergence hypothesis for the clubs and show diverging behavior of the diverging group. The boxplots in Figure 7 reveal the heterogeneity (homogeneity) between (within) the clubs.

Analyzing log t regression scatter-plots for regional data offers similar results to the PWT data. Figure 8 exemplifies the results for final ordering. Most of the convergence clubs offer a parabolic and convex shape which means a nonlinear relationship but no harm for convergence interpretation. But, for the second club the regressand becomes smaller in the last period which rises doubt on the club convergence

Investigating β-convergence yields to the regression output displayed in Table 9. The estimated coefficient is significantly negative. The estimated coefficients of the baseline model (6) briefed in Table 10 offer β-convergence for all clubs, but divergence for the divergence group. The estimated regression lines for the ten convergence clubs after merging are displayed in Figure 9. As the p-values in Table 11) reveal, for ordering rules (I) and (II) the hypothesis of correct parametric specification of the baseline model (6) cannot be rejected at any reasonable significance level, while there are hints for misspecification in (III) and (IV). For assessing the quality of clubbing we estimate nonparametric models for all cases. The estimated bandwidths are displayed in Table 11. With the exception of (II) the bandwidths for $\log(y_{i,0})$ point to nonlinear influences of the regressor. The bandwidths for the club variable are all close to zero. Thus, club compositions are well chosen for all ordering rules. The out-of-sample cross validation analysis offers a strict ranking of ordering rules (I)$\succ$(IV)$\succ$(II)$\succ$(III), where $\succ$ means that the ordering rule on the left has a lower ASEP than the rule on the right.

5.3 Prefecture-level data from Japan

In addition to PWT data on country-level and german regional data on district-level we analyze an in-between — data on 47 Japanese prefectures between 1956 and 1990[8]. The results on merged clubs are displayed in Table 13. Using final ordering and difference ordering we find three convergence clubs which can be merged to two clubs. Average ordering and decreasing weights ordering propose exactly the same results. There are four convergence clubs and one divergence group consisting of three countries. After merging there are only two clubs and one divergence group. The relative transition coefficients over time for the convergence clubs after merging are exemplarily shown for final ordering in Figure 10, where club convergence is indicated as the transition coefficients converge to one. Again, heterogeneity between clubs can be observed from Figure 11. The scatter-plots of log t regression for final ordering displayed in Figure 12 show quite different results than for the other examples. For all three clubs the first half of the time horizon show parabolic and convex points. In the second half, the points of clubs one and three stagnate in contrast to convergence assumption. The results of log t regressions should be handled with care. Analyzing classical β-convergence (1) proposes a positive coefficient which is significant on 10%-level

[8]The data of Barro and Sala-i-Martin (2004) are downloaded from http://www.columbia.edu/∼xs23/data.htm at June 15, 2011.

(see Table 14)[9]. Thus, there are no hints for β-convergence over all prefectures. The estimated convergence coefficients for the baseline model reveal negative signs for both clubs, which are, however, not significantly different from zero (see Table 15). The estimated regression lines are plotted in Figure 13.

Investigating parametric misspecification the HLR test has small p-values (with maximum of around 11%) for all ordering rules. For all ordering rules the bandwidth of $club_i$ is approximately 0.01 or even smaller. Thus, the clusters seems to be well chosen for all methods. The bandwidth for $\log(y_{i,0})$ proposes a linear influence of this variable for ordering rule (I) and (III) and a nonlinear influence for (II) and (IV). The p-values and bandwidths for nonparametric regression can be found in Table 16. The out-of-sample cross-validation offers the following sequence of ordering rules (I)$\succ$(III)$\succ$(II)=(IV).

6 Conclusion

As classical convergence regressions often fail to account for heterogeneity and nonlinearity and recent contributions are able to address either the one or the other, a simple two-step method is proposed to address both issues. Employing a slightly augmented version of the clubbing algorithm of Phillips and Sul (2007a,b, 2009) in step one, we find (i) considerable sensitivity of results on convergence club structures with respect to different initial data orderings. Further, (ii) visual inspections of $\log t$ regression scatter plots reveal that the "convergence interpretation" of the results of such a linear regression should be handled with care. As a second step we propose the use of a nonparametric test and regression which allows to analyze convergence effects on both individual and club level while alleviating potential misclassification in the club formation process using simultaneous smoothing over the club structure.

[9] Again, we only present results for initial ordering.

Three empirical exercises using data on different levels of aggregation, countries from the Penn World Tables, Japanese prefectures, and districts from reunified Germany, respectively, illustrate the proposed two-step approach. For all applications, we find considerable evidence for club-based heterogeneity in convergence analysis by adding the clubs identified in step one as a categorical covariate. Our nonparametric estimation results suggest that the club composition is well chosen. An out-of-sample analysis reveals that initial ordering rule for starting the club identification algorithm (in step one) proposed by Phillips and Sul performs best.

References

[1] Alfò, M., G. Trovato and R.J. Waldmann, (2008). Testing for Country Heterogeneity in Growth Models Using a Finite Mixture Approach. *Journal of Applied Econometrics* **23**, 487-514. DOI: 10.1002/jae.1008

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

[3] Barro, R.J. and X. Sala-i-Martin, (2004). *Economic Growth.* The MIT Press.

[4] Barro, R.J., X. Sala-i-Martin, O.J. Blanchard and R.E. Hall, (1991). Convergence Across States and Regions. *Brookings Papers on Economic Activity.* **1**, 107-182.

[5] Canarella, G. and S. Pollard, (2004). Parameter Heterogeneity in the Classical Growth Model: A Quantile Regression Approach. *Journal of Economic Development* **29**, 1-31. DOI:10.1016/j.econlet.2008.07.011

[6] Canova, F., (2004). Testing for Convergence Clubs in Income per Capita: A Predictive Density Approach. *International Economic Review* **45**: 49-77. DOI: 10.1111/j.1468-2354.2004.00117.x

[7] Durlauf, S.N. and P.A. Johnson, (1995). Multiple Regimes and Cross-Country Growth Behavior. *Journal of Applied Econometrics* **10**: 365-384. DOI: 10.1002/jae.3950100404

[8] Haupt, H. and V. Petring, (2011). Assessing parametric misspecification and heterogeneity in growth regression. *Applied Economics Letters* **18**: 389-394. DOI: 10.1080/13504851003670643

[9] Hayfield, T. and J.S. Racine, (2008). Nonparametric Econometrics: The `np` Package. *Journal of Statistical Software* **27**

[10] Henderson, D.J., (2010). A Test for Multimodality of Regression Derivatives with Application to nonparametric Growth Regression. *Journal of Applied Econometrics*, **25**: 458-480. DOI: 10.1002/jae.1099

[11] Hsiao, C., Q. Li and J.S. Racine, (2007). A Consistent Model Specification Test with Mixed Discrete and Continuous Data. *Journal of Econometrics*, **140**: 802-826. DOI:10.1016/j.jeconom.2006.07.015

[12] Juárez, M.A. and M.F.J. Steel, (2010). Model-Based Clustering of Non-Gaussian Panel Data Besed on Skew-t Distributions. *Journal of Business & Economic Statistics* **28**: 52-66. DOI: 10.1198/jbes.2009.07145

[13] Kalaitzidakis, P., T.P. Mamuneas, A. Savvides and T. Stengos, (2001). Measures of Human Capital and Nonlinearities in Economic Growth. *Journal of Economic Growth* **6**, 229-254. DOI: 10.1023/A:1011347816503

[14] Li, Q. and J.S. Racine, (2007). *Nonparametric Econometrics: Theory and Practice.* Princeton University Press.

[15] Li, Q. and J.S. Racine, (2004). Cross-Validated Local Linear Nonparametric Regression. *Statistica Sinica* **14**: 485-512.

[16] Li, Q. and J.S. Racine, (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business and Economic Statistics* **26**: 423-34. DOI:10.1198/073500107000000250

[17] Liu, Z. and T. Stengos, (1999). Non-Linearities in Cross-Country Growth Regressions: A Semiparametric Approach. *Journal of Applied Econometrics* **14**, 527-538. DOI: 10.1002/(SICI)1099-1255(199909/10)14:5<527::AID-JAE528>3.0.CO;2-X

[18] Mankiw, N.G., D. Romer and D.N. Weil, (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* **107**: 407-437. DOI: 10.2307/2118477

[19] Masanjala, W.H. and C. Papageorgiou, (2004). The Solow Model With CES Technology: Nonlinearities and Parameter Heterogeneity. *Journal of Applied Econometrics* **19**, 171-201. DOI: 10.1002/jae.722

[20] Maasoumi, E., Q. Li and J.S. Racine, (2007). Growth and Convergence: a Profile of Distribution Dynamics and Mobility. *Journal of Econometrics* **136**, 483-508. DOI:10.1016/j.jeconom.2005.11.012

[21] Phillips, P.C.B. and D. Sul, (2003). The elusive empirical shadow of growth convergence. *Cowles Foundation Discussion Paper No. #1398*, Yale University: New Haven CT.

[22] Phillips, P.C.B. and D. Sul, (2007a). Some empirics on economic growth under heterogeneous technology. *Journal of Macroeconomics* **29**: 455-469. DOI:10.1016/j.jmacro.2007.03.002

[23] Phillips, P.C.B. and D. Sul, (2007b). Transition Modeling and Econometric Convergence Tests. *Econometrica* **75**: 6, 1771-1855. DOI: 10.1111/j.1468-0262.2007.00811.x

[24] Phillips, P.C.B. and D. Sul, (2009). Economic Transition and Growth. *Journal of Applied Econometrics* **24**: 1153-1185. DOI: 10.1002/jae.1080

[25] Quah, D., (1993). Empirical cross-section dynamics in economic growth. *European Economic Review*, **37**, 426-434. DOI:10.1016/0014-2921(93)90031-5

[26] Quah, D., (1997). Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs. *Journal of Economic Growth*, **2**, 27-59. DOI: 10.1023/A:1009781613339

[27] Racine, J.S. and Q. Li, (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* **119**: 99-130. DOI: 10.1016/S0304-4076(03)00157-X

[28] Sala-i-Martin, X., (1996). The Classical Approach to Convergence Analysis. *The Economic Journal*, **106**, 1019-1036.

A Tables and Figures

A.1 PWT country-level data

(I) Final ordering				(II) Average ordering			
a) initial classification		b) final classification		a) initial classification		b) final classification	
	$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)
Club 1 [50]	0.38 (0.04)	Club 1 [50]	0.38 (0.04)	Club 1 [67]	0.09 (0.03)	Club 1 [67]	0.09 (0.03)
Club 2 [30]	0.24 (0.03)	Club 2 [30]	0.24 (0.03)	Club 2 [8]	0.36 (0.04)	Club 2 [18]	0.03 (0.03)
Club 3 [21]	0.11 (0.03)	Club 3 [21]	0.11 (0.03)	Club 3 [10]	-0.001 (0.02)		
Club 4 [24]	0.13 (0.06)	Club 4 [38]	-0.44 (0.07)	Club 4 [12]	-0.01 (0.06)	Club 3 [12]	-0.01 (0.06)
Club 5 [14]	0.19 (0.11)			Club 5 [21]	0.03 (0.05)	Club 4 [23]	0.04 (0.05)
Club 6 [11]	1.00 (0.17)	Club 5 [11]	1.00 (0.17)	Club 6 [2]	0.10 (0.31)		
Club 7 [2]	-0.47 (0.84)	Club 6 [2]	-0.47 (0.84)	Club 7 [9]	0.07 (0.05)	Club 5 [16]	0.06 (0.10)
				Club 8 [7]	0.15 (0.12)		
				Club 9 [10]	1.39 (0.15)	Club 6 [10]	1.39 (0.15)
				Group 10 [6]	-2.04^{*} (0.02)	Group 7 [6]	-2.04^{*} (0.02)

(III) Difference ordering				(IV) Decreasing weights ordering			
a) initial classification		b) final classification		a) initial classification		b) final classification	
	$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)
Club 1 [67]	-0.003 (0.007)	Club 1 [67]	-0.003 (0.007)	Club 1 [73]	0.01 (0.03)	Club 1 [73]	0.01 (0.03)
Club 2 [32]	0.71 (0.06)	Club 2 [32]	0.71 (0.06)	Club 2 [24]	0.09 (0.02)	Club 2 [24]	0.09 (0.02)
Club 3 [42]	-0.05 (0.05)	Club 3 [42]	-0.05 (0.05)	Club 3 [22]	0.05 (0.05)	Club 3 [31]	-0.05 (0.05)
Club 4 [4]	1.48 (0.09)	Club 4 [4]	1.48 (0.09)	Club 4 [9]	0.08 (0.06)		
Club 5 [6]	0.43 (0.12)	Club 5 [6]	0.43 (0.12)	Club 5 [2]	0.08 (0.19)	Club 4 [2]	0.08 (0.19)
Group 6 [1]		Group 6 [1]		Club 6 [7]	0.15 (0.11)	Club 5 [15]	-0.07 (0.12)
				Club 7 [8]	1.411 (0.18)		
				Group 8 [7]	-1.80^{*} (0.02)	Group 6 [7]	-1.80^{*} (0.02)

Table 1: Results of clubbing algorithm for PWT data. Club sizes (in brackets), estimates for γ and standard errors of the log t regression (4) are displayed for different ordering rules. a) of each ordering rule gives the initial classification before club merging, b) gives the final classification after merging.

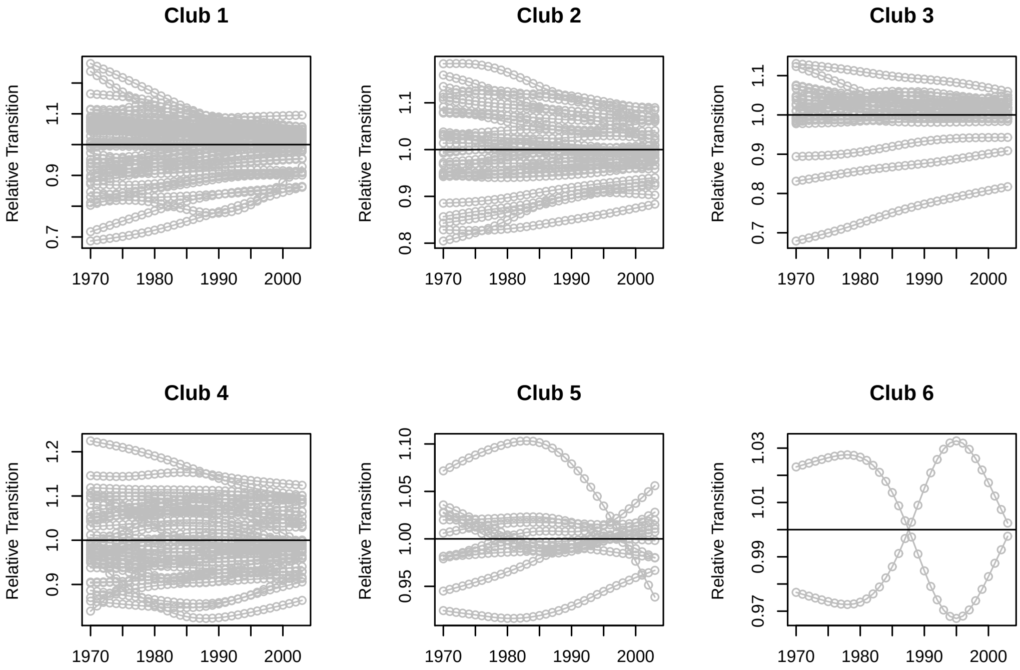

Figure 1: Relative transition coefficients over time for convergence clubs resulting from final ordering after merging, corresponding to (I) b) in Table 1.

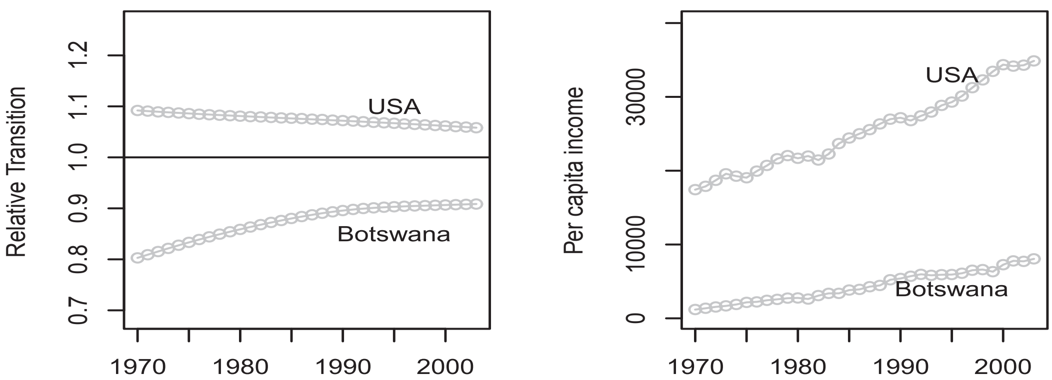

Figure 2: Relative transition coefficients over time and absolute per capita income over time of USA and Botswana.

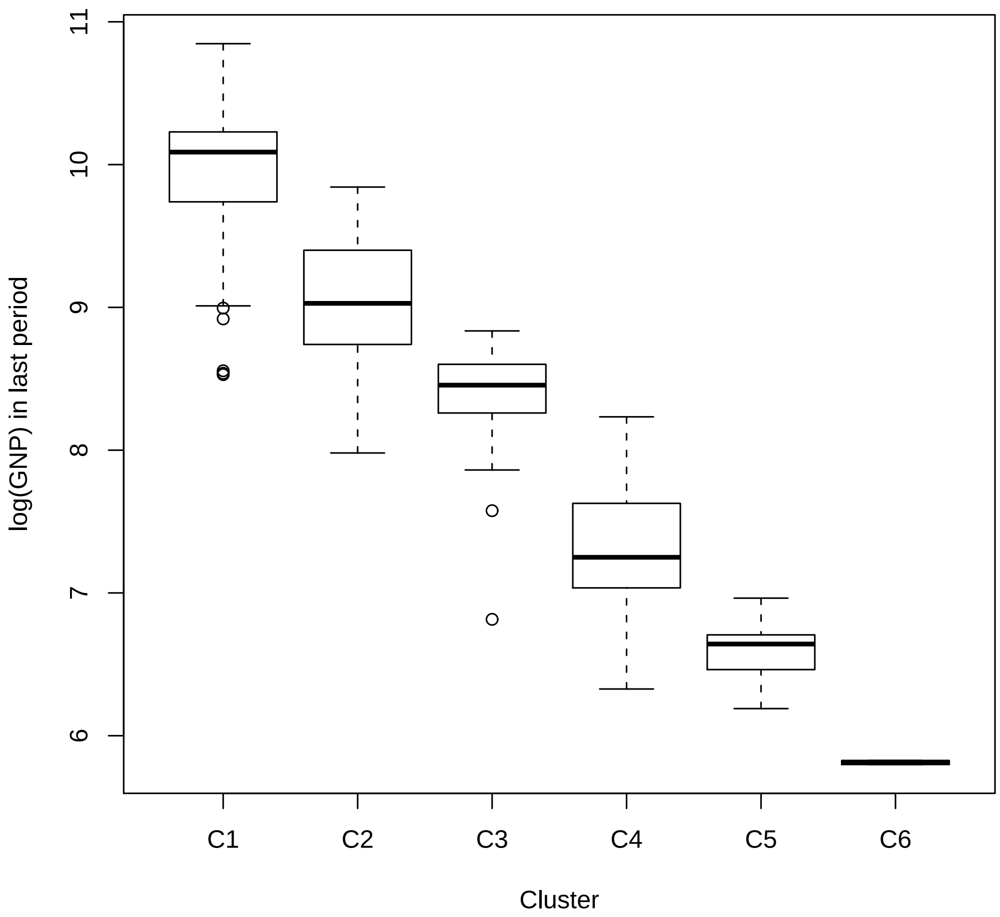

Figure 3: Boxplots of income in final period divided by the convergence clubs resulting from final ordering after merging, corresponding to (I) b) in Table 6.

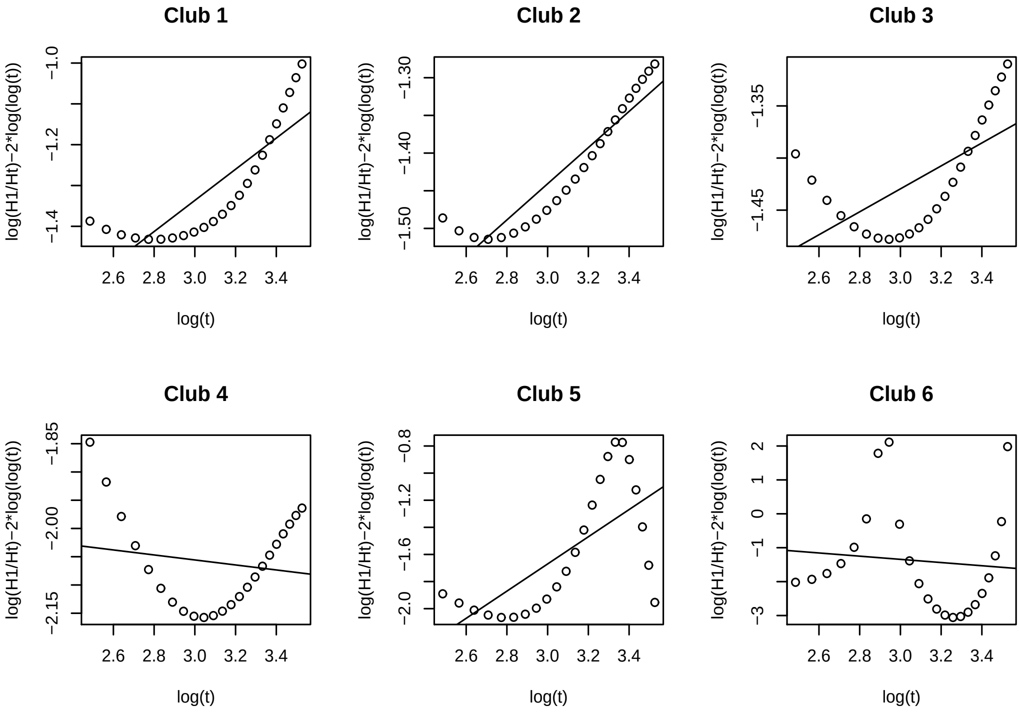

Figure 4: Scatterplots of the log t regressions (4) for clubs found by final ordering after merging, corresponding to (I) b) in Table 1. Solid line is ordinary least squares estimate.

Table 2: Comparison of club number, size, and composition for PWT data and final ordering (I) for different time horizons. The club structure for complete time horizon 1970 to 2003 (partial horizon from 1978 to 2003) is given in rows (columns).

	C1	C2	C3	C4	C5	C6	G7	n_c
C1	49	0	0	0	0	0	1	50
C2	13	16	1	0	0	0	0	30
C3	0	10	8	3	0	0	0	21
C4	0	0	1	23	14	0	0	38
C5	0	0	0	0	4	7	0	11
C6	0	0	0	0	0	0	2	2
n_c	62	26	10	26	18	7	3	152

Table 3: Comparison of club number, size, and composition for PWT data and final ordering (I) for different time horizons. The club structure for complete time horizon 1970 to 2003 (partial horizon from 1986 to 2003) is given in rows (columns).

	C1	C2	C3	C4	C5	C6	C7	C8	G9	n_c
C1	43	4	2	0	0	0	0	1	0	50
C2	5	10	13	2	0	0	0	0	0	30
C3	0	0	9	9	2	1	0	0	0	21
C4	0	0	0	4	8	22	4	0	0	38
C5	0	0	0	0	0	1	10	0	0	11
C6	0	0	0	0	0	0	0	1	1	2
n_c	48	14	24	15	10	24	14	2	1	152

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.5997	0.4037	1.49	0.1395
$\log(y_{i,0})$	-0.0126	0.0495	-0.25	0.7995

$adj.R^2$=-0.006, AIC=-752.24, n=152

Table 4: OLS estimates of classical convergence model (1) for PWT data.

	Estimate	Std. Error	t value	Pr(>|t|)
Club 1	5.9291	0.3352	17.69	0.0000
Club 2	4.1480	0.5241	7.91	0.0000
Club 3	4.0241	0.6178	6.51	0.0000
Club 4	2.9553	0.5387	5.49	0.0000
Club 5	7.1889	2.3127	3.11	0.0023
Group 6	11.3003	8.4093	1.34	0.1812
Club 1:log($y_{i,0}$)	-0.5566	0.0373	-14.92	0.0000
Club 2:log($y_{i,0}$)	-0.4191	0.0620	-6.75	0.0000
Club 3:log($y_{i,0}$)	-0.4499	0.0780	-5.77	0.0000
Club 4:log($y_{i,0}$)	-0.3899	0.0750	-5.20	0.0000
Club 5:log($y_{i,0}$)	-1.0743	0.3279	-3.28	0.0013
Club 6:log($y_{i,0}$)	-1.7197	1.1323	-1.52	0.1311

$adj.R^2$=0.8486, AIC=-1014.564, N=152

Table 5: OLS estimates of baseline model (7) for PWT data.

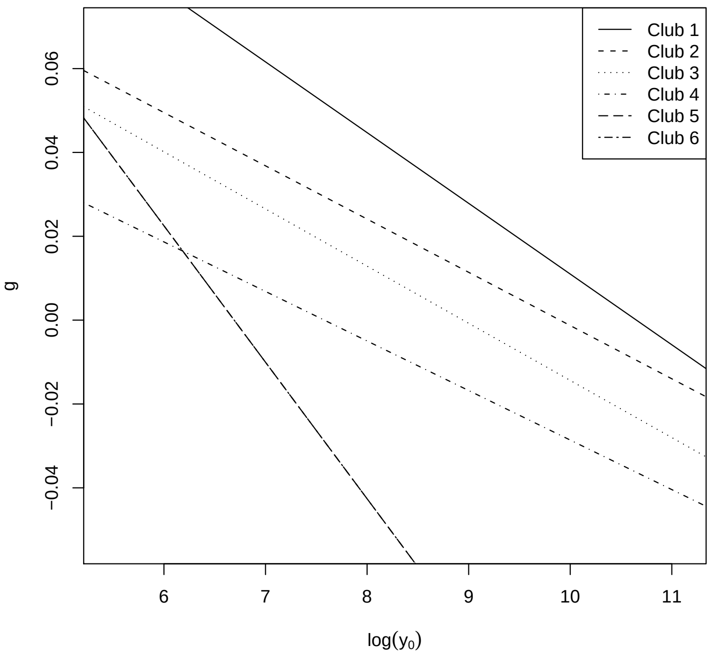

Figure 5: Estimated regression lines from the estimation displayed in Table 3 for the five convergence clubs for PWT data.

	bandwidth of $\log(y_{i,0})$	bandwidth of *club*	p-value of Hsiao-Li-Racine test
(I)	1.054	0.006	0.0125
(II)	0.839	0.01	≈ 0
(III)	0.812	≈ 0	≈ 0
(IV)	1.1205	0.028	≈ 0

Table 6: Estimated bandwidths for nonparametric baseline model estimation using a mixed kernel estimation for PWT data and ordering rules (I)-(IV) and p-values for Hsiao-Li-Racine tests.

	(I)	(II)	(III)	(IV)
(I)	–	0.17	0.21	0.07
(II)	0.83	–	0.49	0.25
(III)	0.79	0.51	–	0.34
(IV)	0.93	0.75	0.66	–

Table 7: Pairwise comparisons of cross-validation performance. Number equals share of $B = 10,000$ replications in which model in column has smaller ASEP (average squared error of prediction) than model in row for PWT data.

A.2 German district-level data

(I) Final ordering				(II) Average ordering			
a) initial classification		b) final classification		a) initial classification		b) final classification	
	$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)
Club 1 [3]	0.84 (0.26)	Club 1 [3]	0.84 (0.26)	Club 1 [3]	0.84 (0.26)	Club 1 [3]	0.84 (0.26)
Club 2 [5]	0.01 (0.05)	Club 2 [5]	0.01 (0.05)	Club 2 [3]	0.99 (0.30)	Club 2 [3]	0.99 (0.30)
Club 3 [4]	0.19 (0.16)	Club 3 [10]	0.03 (0.14)	Club 3 [4]	0.34 (0.16)	Club 3 [12]	-0.08 (0.12)
Club 4 [6]	0.15 (0.16)			Club 4 [8]	0.11 (0.15)		
Club 5 [30]	0.09 (0.14)	Club 4 [30]	0.09 (0.14)	Club 5 [20]	0.08 (0.14)	Club 4 [33]	0.02 (0.13)
Club 6 [24]	0.15 (0.16)	Club 5 [72]	-0.14 (0.11)	Club 6 [13]	0.13 (0.16)		
Club 7 [14]	0.12 (0.16)			Club 7 [26]	0.08 (0.15)	Club 5 [48]	0.01 (0.13)
Club 8 [14]	0.11 (0.15)			Club 8 [8]	0.05 (0.15)		
Club 9 [7]	0.03 (0.14)			Club 9 [14]	0.11 (0.15)		
Club 10 [13]	0.30 (0.15)			Club 10 [33]	0.21 (0.15)	Club 6 [86]	-0.07 (0.12)
Club 11 [16]	0.20 (0.16)	Club 6 [76]	-0.07 (0.12)	Club 11 [6]	1.42 (0.19)		
Club 12 [33]	0.11 (0.15)			Club 12 [17]	0.39 (0.17)		
Club 13 [27]	0.16 (0.16)			Club 13 [12]	0.07 (0.15)		
Club 14 [90]	0.10 (0.15)	Club 7 [90]	0.01 (0.15)	Club 14 [18]	0.11 (0.16)		
Club 15 [80]	0.15 (0.14)	Club 8 [80]	0.15 (0.14)	Club 15 [85]	-0.05 (0.13)	Club 7 [134]	-0.13 (0.11)
Club 16 [56]	0.04 (0.11)	Club 9 [56]	0.04 (0.11)	Club 16 [2]	0.66 (1.97)		
Club 17 [13]	0.09 (0.12)	Club 10 [13]	0.09 (0.12)	Club 17 [30]	0.06 (0.15)		
Group 18 [4]	-1.33* (0.03)	Group 11 [4]	-1.33* (0.03)	Club 18 [17]	0.05 (0.15)		
				Club 19 [18]	0.04 (0.15)	Club 8 [85]	-0.18 (0.11)
				Club 20 [26]	0.04 (0.14)		
				Club 21 [25]	0.03 (0.14)		
				Club 22 [16]	0.11 (0.15)		
				Club 23 [16]	0.56 (0.17)	Club 9 [16]	0.56 (0.17)
				Club 24 [8]	-0.03 (0.12)	Club 10 [8]	-0.03 (0.12)
				Group 25 [11]	-1.39* (0.02)	Group 11 [11]	-1.39* (0.02)

(III) Difference ordering				(IV) Decreasing Weights ordering			
a) initial classification		b) final classification		a) initial classification		b) final classification	
	$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)		$\hat{\gamma}$ (SE of $\hat{\gamma}$)
Club 1 [67]	-0.01 (0.01)	Club 1 [114]	-0.11 (0.09)	Club 1 [5]	0.31 (0.09)	Club 1 [5]	0.31 (0.09)
Club 2 [32]	0.66 (0.07)			Club 2 [8]	0.11 (0.15)	Club 2 [8]	0.11 (0.15)
Club 3 [41]	0.67 (0.15)	Club 2 [83]	-0.10 (0.07)	Club 3 [24]	0.03 (0.14)	Club 3 [32]	0.14 (0.37)
Club 4 [11]	0.36 (0.18)			Club 4 [8]	0.18 (0.16)		
Club 5 [27]	0.50 (0.09)			Club 5 [26]	0.08 (0.15)	Club 4 [48]	0.01 (0.13)
Club 6 [34]	-0.07 (0.05)	Club 3 [34]	-0.07 (0.05)	Club 6 [8]	0.05 (0.15)		
Club 7 [24]	0.35 (0.67)	Club 4 [24]	0.35 (0.67)	Club 7 [14]	0.11 (0.15)		
Club 8 [36]	-0.09 (0.12)	Club 5 [36]	-0.09 (0.12)	Club 8 [32]	0.21 (0.15)	Club 5 [94]	-0.14 (0.11)
Group 9 [148]	-1.08* (0.04)	Group 6 [148]	-1.08* (0.04)	Club 9 [14]	0.21 (0.16)		
				Club 10 [4]	1.21 (0.32)		
				Club 11 [4]	0.48 (0.12)		
				Club 12 [6]	0.34 (0.17)		
				Club 13 [12]	0.07 (0.15)		
				Club 14 [22]	-0.01 (0.14)		
				Club 15 [46]	0.38 (0.14)	Club 6 [141]	-0.17 (0.11)
				Club 16 [3]	0.52 (0.83)		
				Club 17 [81]	-0.63 (0.13)		
				Club 18 [3]	3.16 (0.96)		
				Club 19 [5]	2.52 (0.58)		
				Club 20 [3]	0.82 (0.17)		
				Club 21 [53]	-0.03 (0.11)	Club 7 [72]	-0.13 (0.11)
				Club 22 [15]	0.94 (0.25)		
				Club 23 [4]	0.42 (0.24)		
				Club 24 [22]	-0.10 (0.08)	Club 8 [22]	-0.10 (0.08)
				Group 25 [17]	-1.28* (0.03)	Group 9 [17]	-1.28* (0.03)

Table 8: Results of clubbing algorithm for German district data. Club sizes (in brackets), estimates for γ and standard errors of the log t regression (4) are displayed for different ordering rules. a) of each ordering rule gives the initial classification before club merging, b) gives the final classification after merging.

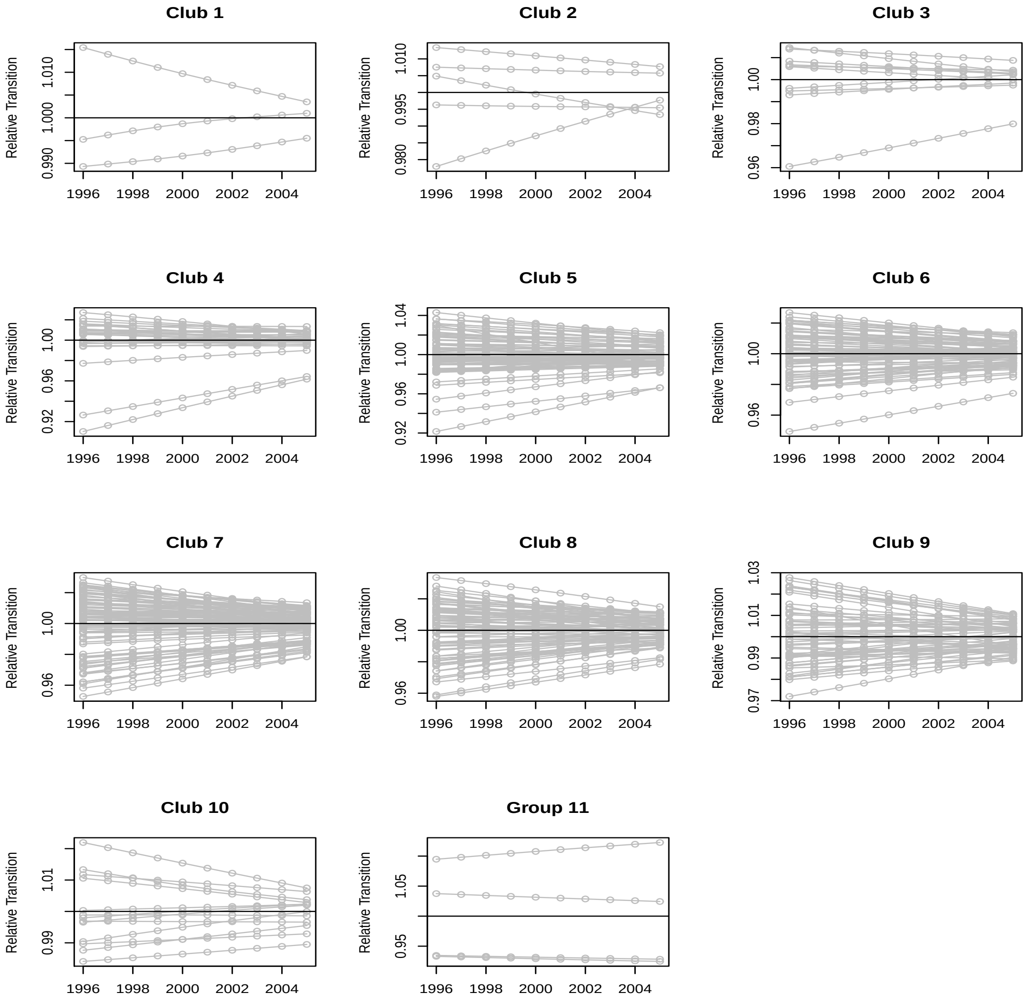

Figure 6: Relative transition coefficients over time for convergence clubs resulting from final ordering after merging, corresponding to (I) b) in Table 6.

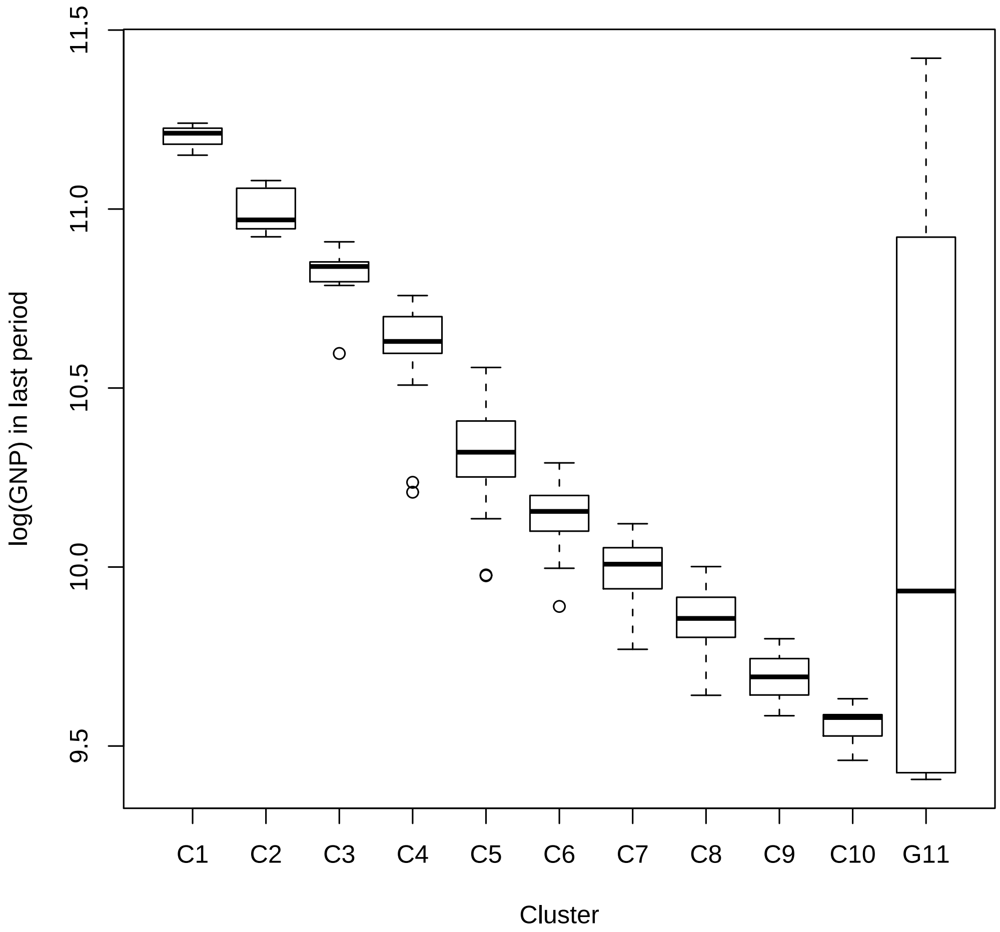

Figure 7: Boxplots of income in final period divided by the convergence clubs resulting from final ordering after merging, corresponding to (I) b) in Table 6.

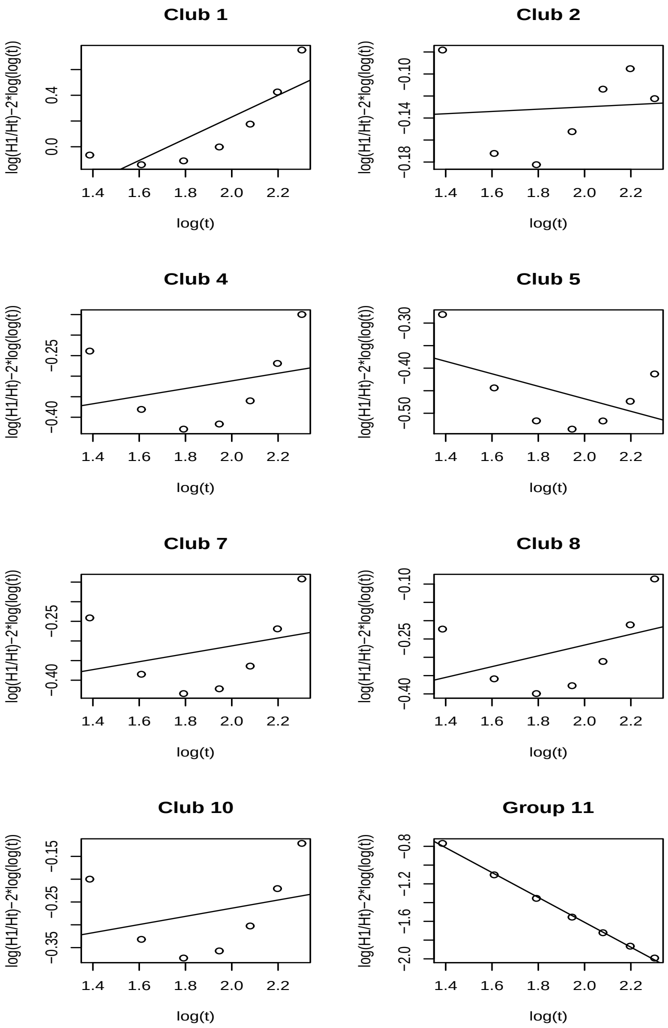

Figure 8: Scatterplots of log t regression for clubs found by final ordering after merging, corresponding to (I) b) in Table 6. Solid line is ordinary least squares estimate.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.9115	0.1337	6.82	0.0000
$\log(y_{i,0})$	-0.0734	0.0135	-5.43	0.0000

$adj.R^2$=0.0632, AIC=-790.441, $N = 439$

Table 9: OLS estimates of classical convergence model (1) for German district data.

	Estimate	Std. Error	t value	Pr(>|t|)
Club 1	10.0508	1.3900	7.23	0.0000
Club 2	8.6208	1.2689	6.79	0.0000
Club 3	5.4437	0.9146	5.95	0.0000
Club 4	5.2388	0.3163	16.57	0.0000
Club 5	4.8481	0.2306	21.02	0.0000
Club 6	5.2649	0.3290	16.00	0.0000
Club 7	5.4732	0.2214	24.72	0.0000
Club 8	5.5146	0.2706	20.38	0.0000
Club 9	5.4455	0.3927	13.87	0.0000
Club 10	5.2947	0.9082	5.83	0.0000
Group 11	-1.5484	0.2915	-5.31	0.0000
Club 1:log($y_{i,0}$)	-0.8943	0.1284	-6.96	0.0000
Club 2:log($y_{i,0}$)	-0.7807	0.1180	-6.62	0.0000
Club 3:log($y_{i,0}$)	-0.4938	0.0862	-5.73	0.0000
Club 4:log($y_{i,0}$)	-0.4851	0.0303	-16.00	0.0000
Club 5:log($y_{i,0}$)	-0.4589	0.0228	-20.11	0.0000
Club 6:log($y_{i,0}$)	-0.5093	0.0330	-15.41	0.0000
Club 7:log($y_{i,0}$)	-0.5397	0.0226	-23.88	0.0000
Club 8:log($y_{i,0}$)	-0.5519	0.0280	-19.74	0.0000
Club 9:log($y_{i,0}$)	-0.5546	0.0412	-13.46	0.0000
Club 10:log($y_{i,0}$)	-0.5499	0.0960	-5.73	0.0000
Group 11:log($y_{i,0}$)	0.1780	0.0293	6.07	0.0000

$adj.R^2$=0.852, AIC=-1567.249, $N = 439$

Table 10: OLS estimates of baseline model (7) for German district data.

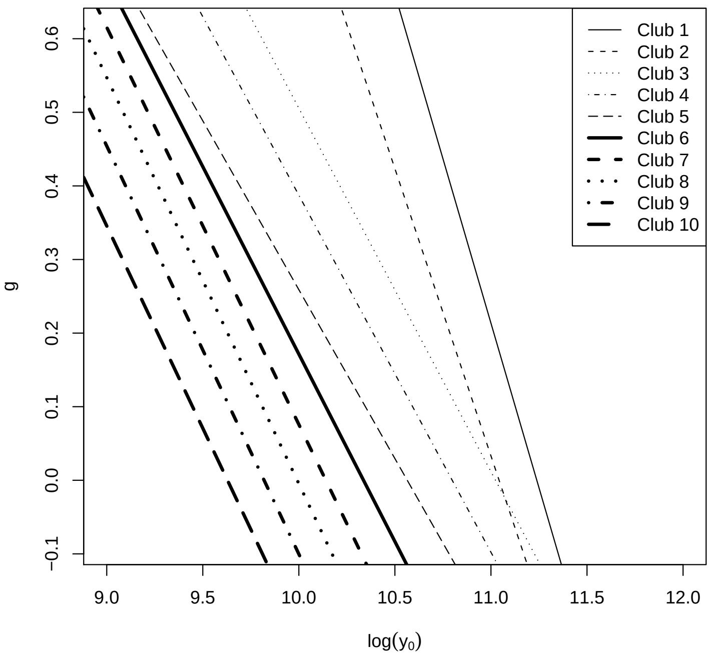

Figure 9: Estimated regression lines from the estimation displayed in Table 8 for the ten convergence clubs for German district data.

	bandwidth of $\log(y_{i,0})$	bandwidth of $club$	p-value of Hsiao-Li-Racine test
(I)	0.133	0.0002	0.61
(II)	16092882	0.007	0.99
(III)	0.121	0.0013	0.01
(IV)	0.176	0.0053	≈ 0

Table 11: Estimated bandwidths for nonparametric baseline model estimation using a mixed kernel estimation for German district data and ordering rules (I)-(IV) and p-values for Hsiao-Li-Racine tests.

	(I)	(II)	(III)	(IV)
(I)	–	0.24	0.18	0.25
(II)	0.76	–	0.16	0.56
(III)	0.82	0.84	–	0.89
(IV)	0.75	0.44	0.11	–

Table 12: Pairwise comparisons of cross-validation performance. Number equals share of $B = 10,000$ replications in which model in column has smaller ASEP (average squared error of prediction) than model in row for regional data.

A.3 Japanese prefecture-level data

(I) Final ordering				**(II) Average ordering**			
a) initial classification		**b) final classification**		**a) initial classification**		**b) final classification**	
	$\hat{\gamma}$ (SE of $\hat{\gamma}$)		**$\hat{\gamma}$ (SE of $\hat{\gamma}$)**		**$\hat{\gamma}$ (SE of $\hat{\gamma}$)**		**$\hat{\gamma}$ (SE of $\hat{\gamma}$)**
Club 1 [28]	0.09 (0.01)	Club 1 [28]	0.09 (0.01)	Club 1 [17]	0.12 (0.01)	Club 1 [35]	0.01 (0.01)
Club 2 [17]	0.09 (0.01)	Club 2 [19]	0.00 (0.01)	Club 2 [10]	0.10 (0.01)		
Club 3 [2]	0.10 (0.02)			Club 3 [8]	0.05 (0.01)		
				Club 4 [9]	0.18 (0.00)	Club 2 [9]	0.18 (0.00)
				Group 5 [3]	-0.47 (0.01)	Group 3 [3]	-0.47 (0.01)

(III) Difference ordering				**(IV) Decreasing Weights ordering**			
a) initial classification		**b) final classification**		**a) initial classification**		**b) final classification**	
	$\hat{\gamma}$ (SE of $\hat{\gamma}$)		**$\hat{\gamma}$ (SE of $\hat{\gamma}$)**		**$\hat{\gamma}$ (SE of $\hat{\gamma}$)**		**$\hat{\gamma}$ (SE of $\hat{\gamma}$)**
Club 1 [12]	0.23 (0.01)	Club 1 [29]	0.08 (0.01)	Club 1 [17]	0.12 (0.01)	Club 1 [35]	0.01 (0.01)
Club 2 [17]	0.22 (0.01)			Club 2 [10]	0.10 (0.01)		
Club 3 [18]	0.02 (0.01)	Club 2 [18]	0.02 (0.01)	Club 3 [8]	0.05 (0.01)		
				Club 4 [9]	0.18 (0.00)	Club 2 [9]	0.18 (0.00)
				Group 5 [3]	-0.47 (0.01)	Group 3 [3]	-0.47 (0.01)

Table 13: Results of clubbing algorithm for Japanese prefecture-level data. Club sizes (in brackets), estimates for γ and standard errors of the log t regression (4) are displayed for different ordering rules. a) of each ordering rule gives the initial classification before club merging, b) gives the final classification after merging.

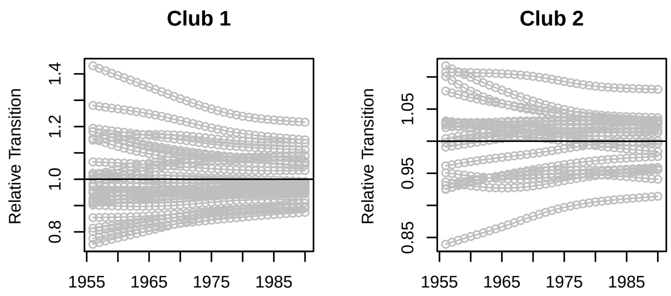

Figure 10: Relative transition coefficients over time for convergence clubs resulting from final ordering after merging, corresponding to (I) b) in Table 11.

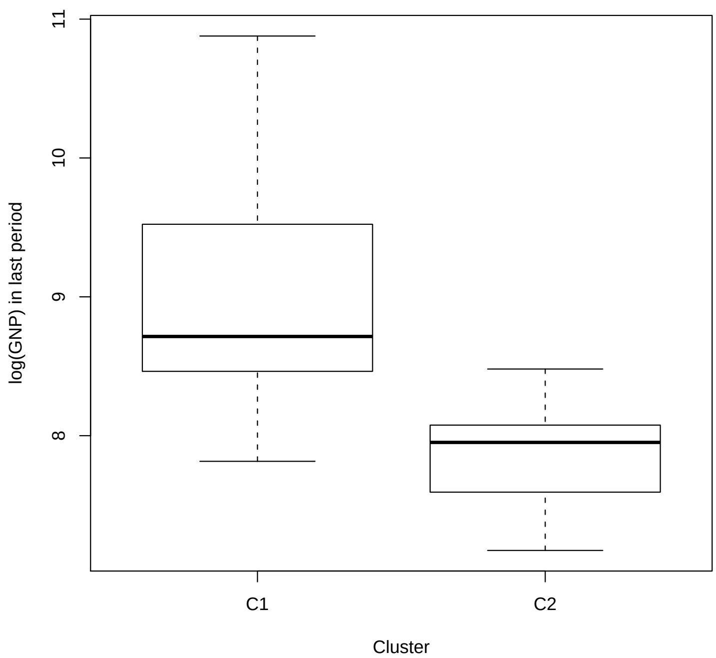

Figure 11: Boxplots of income in final period divided by the convergence clubs resulting from final ordering after merging, corresponding to (I) b) in Table 6.

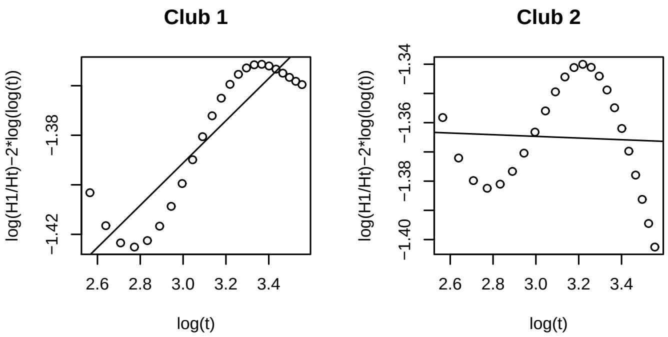

Figure 12: Scatterplots of log t regressions for clubs found by final ordering after merging, corresponding to (I) b) in Table 11. Solid line is ordinary least squares estimate.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.1863	0.2885	11.05	0.0000
$\log(y_{i,0})$	0.1037	0.0597	1.74	0.0891
$adj.R^2$=0.042, AIC=21.17, $N = 47$				

Table 14: OLS estimates of classical convergence model (1) for Japanese prefecture-level data.

	Estimate	Std. Error	t value	Pr(>|t|)
factor(daten)1	3.9644	0.3179	12.47	0.0000
factor(daten)2	3.6674	0.8260	4.44	0.0001
factor(daten)1:bip	-0.0273	0.0619	-0.44	0.6615
factor(daten)2:bip	-0.0453	0.1898	-0.24	0.8124
$adj.R^2$=0.2615, AIC=7.5555, $N = 47$				

Table 15: OLS estimates of baseline model (7) for Japanese prefecture-level data.

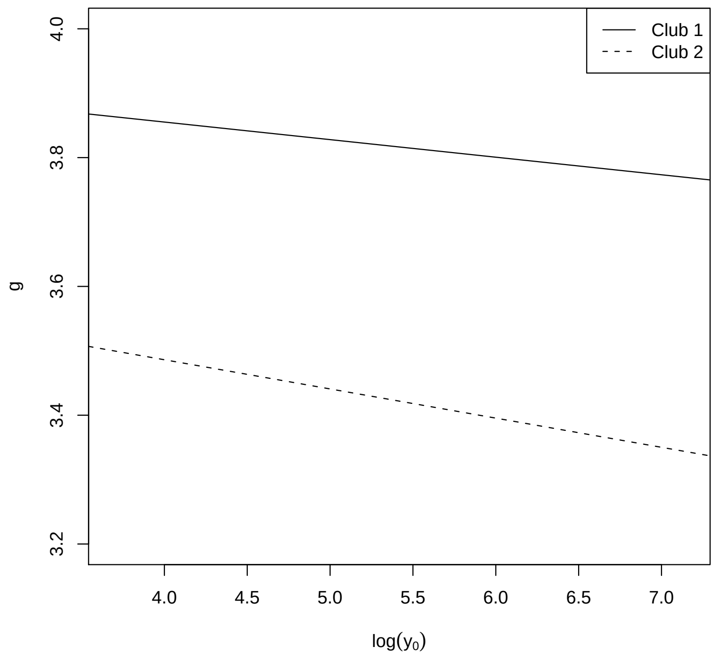

Figure 13: Estimated regression lines from the estimation displayed in Table 13 for the two convergence clubs for Japanese prefecture-level data.

	bandwidth of $\log(y_{i,0})$	bandwidth of *club*	p-value of Hsiao-Li-Racine test
(I)	2721352	0.01	0.04
(II)/(IV)[10]	2.366	0.01	0.11
(III)	1776859	0.003	0.05

Table 16: Estimated bandwidths for nonparametric baseline model estimation using a mixed kernel estimation for Japanese prefecture-level data and ordering rules (I)-(IV) and p-values for Hsiao-Li-Racine tests.

	(I)	(II)/(IV)	(III)
(I)	–	0.43	0.48
(II)/(IV)	0.57	–	0.57
(III)	0.52	0.43	–

Table 17: Pairwise comparisons of cross-validation performance. Number equals share of $B = 10,000$ replications in which model in column has smaller ASEP (average squared error of prediction) than model in row for Japanese prefecture-level data.

A.4 Lists of countries/districts/prefectures in merged clubs using final ordering

Table 18: PWT country level data

Club 1: Antigua, Australia, Austria, Belgium, Bermuda, Botswana, Brunei, Canada, Cape Verde, Chile, China, Cyprus, Denmark, Dominica, Equatorial Guinea, Finland, France, Germany, Hong Kong, Iceland, Ireland, Israel, Italy, Japan, Republic of Korea, Kuwait, Luxembourg, Macao, Malaysia, Maldives, Malta, Mauritius, Netherlands, New Zealand, Norway, Oman, Portugal, Puerto Rico, Qatar, Singapore, Spain, St. Kitts and Nevis, St. Vincent and the Grenadines, Sweden, Switzerland, Taiwan, Thailand United Arab Emirates, United Kingdom, United States
Club 2: Argentina, Bahamas, Bahrain, Barbados, Belize, Brazil, Colombia, Costa Rica, Dominican Republic, Egypt, Gabon, Greece, Grenada, Hungary, India, Indonesia, Mexico, Netherlands Antilles, Panama, Poland, Saudi Arabia, South Africa, Sri Lanka, St. Lucia, Swaziland, Tonga, Trinidad and Tobago, Tunisia, Turkey, Uruguay
Club 3: Algeria, Bhutan, Cuba, Ecuador, El Salvador, Fiji, Guatemala, Iran, Jamaica, Lesotho, Federated States of Micronesia, Morocco, Namibia, Pakistan, Papua New Guinea, Paraguay, Peru, Philippines, Romania, Suriname, Venezuela
Club 4: Benin, Bolivia, Burkina Faso, Cameroon, Cote d'Ivoire, Ethiopia, Ghana, Guinea, Honduras, Jordan, Democratic Republic of Korea, Laos, Mali, Mauritania, Mozambique, Nepal, Nicaragua, Samoa, Solomon Islands, Syria, Tanzania, Uganda, Vanuatu, Zimbabwe, Cambodia, Chad, Comoros, Republic of Congo, The Gambia, Iraq, Kenya, Kiribati, Malawi, Mongolia, Nigeria, Sao Tome and Principe, Senegal, Sudan
Club 5: Afghanistan, Burundi, Central African Republic, Guinea Bissau, Madagascar, Niger, Rwanda, Sierra Leone, Somalia, Togo, Zambia
Club 6: Democratic Republic of Congo, Liberia

Table 19: German district level data

Club 1: Wolfsburg(DFC)[11], Frankfurt am Main(DFC), Schweinfurt(DFC)	
Club 2: Düsseldorf(DFC), Ludwigshafen am Rhein(DFC),Stuttgart(DFC), Ingolstadt(DFC), Regensburg(DFC)	
Club 3: Mannheim(DFC), München(DFC), Erlangen(DFC), Aschaffenburg(DFC), Darmstadt(DFC), Koblenz(DFC), Ulm(DFC),Passau(DFC), Dingolfing-Landau, Bamberg(DFC)	
Club 4: Hamburg(DFC), Bremen(DFC), Köln(DFC), Leverkusen(DFC), Münster(DFC), Offenbach am Main(DFC), Wiesbaden(DFC), Kassel(DFC), Mainz(DFC), Heilbronn(DFC), Baden.Baden(DFC), Karlsruhe(DFC), Heidelberg(DFC), Altötting, Freising, Landshut(DFC), Straubing(DFC), Amberg(DFC), Weiden in der Oberpfalz(DFC), Bayreuth(DFC), Coburg(DFC), Ansbach(DFC), Fürth(DFC), Nürnberg(DFC), Würzburg(DFC), Augsburg(DFC), Kempten im Allgäu(DFC), Memmingen(DFC), Teltow-Fläming, Merseburg-Querfurt	
Club 5: Braunschweig(DFC), Salzgitter(DFC), Emden(DFC), Oldenburg(DFC), Osnabrück(DFC), Essen(DFC), Krefeld(DFC), Rhein-Kreis Neuss, Bonn(DFC), Hochtaunuskreis, Main-Taunus-Kreis, Trier(DFC), Main-Taunus-Kreis, Kaiserslautern(DFC), Landau in der Pfalz(DFC), Freiburg im Breisgau(DFC), Biberach, Rosenheim(DFC), Hof(DFC), Schwabach(DFC), Donau-Ries, Wismar(DFC), Dresden(DFC), Jena(DFC), Sömmerda, Kiel(DFC), Vechta, Aachen(DFC), Dortmund(DFC), Offenbach, Böblingen, Rastatt, Bodenseekreis, Ravensburg, Günzburg, Saarbrücken(DFC), Saarpfalz-Kreis, Zwickau(DFC), Eisenach(DFC), Flensburg(DFC), Bremerhaven(DFC), Fulda, Speyer(DFC), Heilbronn, Pforzheim(DFC), Ortenaukreis, Rottweil, Tuttlingen, Pfaffenhofen an der Ilm, Starnberg, Weilheim-Schongau, Cottbus(DFC), Schwerin(DFC), Region Hannover, Wesermarsch, Hersfeld-Rotenburg, Pirmasens(DFC), Hohenlohekreis, Ostalbkreis, Frankfurt Oder(DFC), Wilhelmshaven(DFC), Bielefeld(DFC), Olpe, Zweibrücken, Ludwigsburg, Schwäbisch Hall, Reutlingen, Mühldorf am Inn, Lichtenfels, Main-Spessart, Neu-Ulm, Potsdam(DFC), Ohrekreis	
Club 6: Duisburg(DFC), Gütersloh, Bochum(DFC), Lahn-Dill-Kreis, Waldeck-Frankenberg, Esslingen, Bad-Tölz-Wolfratshausen, Rosenheim, Deggendorf, Cham, Neumarkt in der Oberpfalz, Kaufbeuren(DFC), Augsburg, Neubrandenburg(DFC), Dessau(DFC), Erfurt(DFC), Lübeck(DFC), Neumünster(DFC), Pinneberg, Stormarn, Osterode am Harz, Stade, Emsland, Mülheim an der Ruhr(DFC), Remscheid.(DFC), Wuppertal(DFC), Herford, Hagen(DFC), Siegen-Wittgenstein, Bergstraße, Main-Kinzig-Kreis, Gießen, Marburg-Biedenkopf, Main-Tauber-Kreis, Karlsruhe, Schwarzwald-Baar-Kreis, Konstanz, Dachau, Neuburg-Schrobenhausen, Traunstein, Schwandorf, Kronach, Ansbach, Erlangen-Höchstadt, Rhön-Grabfeld, Miltenberg, Ostallgäu, Chemnitz(DFC), Magdeburg(DFC), Nordfriesland, Mönchengladbach(DFC), Mettmann, Minden-Lübbecke, Märkischer Kreis, Wetteraukreis, Frankenthal (Pfalz, DFC), Worms(DFC), Germersheim, Rems-Murr-Kreis, Heidenheim, Freudenstadt, Lörrach, Zollernalbkreis, Miesbach, Landshut, Weißenburg-Gunzenhausen, Aschaffenburg, Kitzingen, Lindau(Bodensee), Oberallgäu, Saarlouis, Greifswald(DFC), Rostock(DFC), Leipzig(DFC), Bitterfeld, Suhl(DFC)	

[11]DFC=District-free city

Club 7: Dithmarschen, Segeberg, Steinburg, Göttingen, Diepholz, Hameln-Pyrmont, Hildesheim, Soltau-Fallingbostel, Verden, Cloppenburg, Oberhausen(DFC), Solingen(DFC), Kleve, Rhein-Erft-Kreis, Euskirchen, Oberbergischer-Kreis, Gelsenkirchen(DFC), Borken, Warendorf, Lippe, Paderborn, Ennepe-Ruhr-Kreis, Hochsauerlandkreis, Soest, Odenwaldkreis, Limburg-Weilburg, Kassel, Schwalm-Eder-Kreis, Werra-Meißner-Kreis, Mayen-Koblenz, Neuwied, Rhein-Hunsrück-Kreis, Westerwaldkreis, Bernkastel-Wittlich, Eifelkreis-Bitburg-Prüm, Neustadt an der Weinstraße, Mainz-Bingen, Göppingen, Neckar-Odenwald-Kreis, Rhein-Neckar-Kreis, Enzkreis, Waldshut, Tübingen, Alb-Donau-Kreis, Sigmaringen, Berchtesgadener-Land, Ebersberg, Eichstätt, Erding, Garmisch-Partenkirchen, Landsberg am Lech, Kelheim, Passau, Regen, Rottal-Inn, Tirschenreuth, Hof, Kulmbach, Wunsiedel im Fichtelgebirge, Nürnberger-Land, Roth, Bad-Kissingen, Haßberge, Würzburg, Aichach-Friedberg, Dillingen an der Donau, Unterallgäu, St. Wendel, Berlin(DFC), Brandenburg an der Havel, Oder-Spree, Uckermark, Stralsund(DFC), Annaberg, Chemnitzer Land, Freiberg, Riesa-Großenhain, Döbeln, Bernburg, Halle (Saale, DFC), Aschersleben-Staßfurt, Jerichower Land, Wernigerode, Altmarkkreis Salzwedel, Gera(DFC), Wartburgkreis, Schmalkalden-Meiningen, Gotha, Sonneberg, Saale-Orla
Club 8: Ostholstein, Rendsburg-Eckernförde, Goslar, Northeim, Holzminden, Nienburg Weser, Celle, Lüchow-Dannenberg, Lüneburg, Rotenburg Wümme, Uelzen, Delmenhorst(DFC,)Ammerland, Friesland, Grafschaft Bentheim, Leer, Osnabrück, Viersen, Wesel, Aachen, Düren, Rheinisch-Bergischer Kreis, Rhein-Sieg-Kreis, Coesfeld, Recklinghausen, Steinfurt, Höxter, Hamm(DFC), Unna, Darmstadt-Dieburg, Rheingau-Taunus-Kreis, Vogelsbergkreis, Altenkirchen (Westerwald), Bad Kreuznach, Birkenfeld, Cochem Zell, Rhein-Lahn-Kreis, Vulkaneifel, Alzey-Worms, Donnersbergkreis, Südliche Weinstraße, Calw, Breisgau Hochschwarzwald, Emmendingen, Fürstenfeldbruck, Freyung-Grafenau, Straubing-Bogen, Amberg-Sulzbach, Neustadt an der Waldnaab, Regensburg, Bamberg, Coburg, Forchheim, Fürth, Neustadt an der Aisch-Bad Windsheim, Merzig-Wadern, Neunkirchen, Dahme (Spreewald), Oberhavel, Oberspreewald-Lausitz, Ostprignitz-Ruppin, Prignitz, Demmin, Müritz, Rügen, Plauen(DFC), Mittweida, Stollberg, Bautzen, Meißen, Kamenz, Torgau-Oschatz, Wittenberg, Weißenfels, Bördekreis, Weimar, Eichsfeld, Hildburghausen, Ilm-Kreis, Saalfeld-Rudolstadt
Club 9: Herzogtum Lauenburg, Plön, Schleswig-Flensburg, Helmstedt, Peine, Schaumburg, Cuxhaven, Harburg, Osterholz, Aurich, Oldenburg, Wittmund, Heinsberg, Bottrop(DFC), Herne(DFC), Ahrweiler, Trier-Saarburg, Bad Dürkheim, Kaiserslautern, Kusel, Bayreuth, Schweinfurt, Elbe-Elster, Potsdam-Mittelmark, Spree-Neiße, Bad Doberan, Güstrow, Ludwigslust, Parchim, Vogtlandkreis, Mittlerer Erzgebirgskreis, Aue-Schwarzenberg, Görlitz, Hoyerswerda, Niederschlesischer-Oberlausitzkreis, Löbau-Zittau, Sächsische Schweiz, Weißeritzkreis, Delitzsch, Muldentalkreis, Anhalt-Zerbst, Köthen, Burgenlandkreis, Mansfelder Land, Saalkreis, Sangerhausen, Halberstadt, Stendal, Quedlinburg, Schönebeck, Nordhausen, Unstrut-Hainich-Kreis, Weimarer Land, Saale-Holzland-Kreis, Greiz, Altenburger Land
Club 10: Gifhorn, Wolfenbüttel, Barnim, Havelland, Märkisch Oderland, Mecklenburg-Strelitz, Nordvorpommern, Nordwestmecklenburg, Ostvorpommern, Uecker-Randow, Zwickauer Land, Leipziger Land, Kyffhäuserkreis

Group 11: Groß Gerau, Rhein-Pfalz-Kreis, Südwestpfalz, München

Table 20: Japanese prefecture level data

Club 1: Hokkaido, Miyagi, Fukushima, Niigata, Ibaraki, Tochigi, Gumma, Saitama, Chiba, Tokyo, Kanagawa, Yamanashi, Nagano, Shizuoka, Gifu, Aichi, Mie, Shiga, Kyoto, Osaka, Hyogo, Nara, Hiroshima, Fukuoka, Kumamoto, Oita, Kagoshima, Okinawa
Club 2: Aomori, Iwate, Akita, Yamagata, Toyama, Ishikawa, Fukui, Tottori, Shimane, Okayama, Yamaguchi, Tokushima, Kagawa, Ehime, Saga, Nagasaki, Miyazaki, Wakayama, Kochi