

Detecting Convergence Clubs*

Fuat C. Beylunioglu[†] Thanasis Stengos[‡] M. Ege Yazgan[§]

April 2018

To be appeared in Macroeconomic Dynamics

Abstract

The convergence hypothesis, which is developed in the context of growth economics, asserts that the income differences across countries are transitory, and developing countries will eventually attain the level of income of developed ones. On the other hand convergence clubs hypothesis claim that the convergence can only be realized across groups of countries that share some common characteristics. In this study, we propose a new method to find convergence clubs that combines a pairwise method of testing convergence with maximum clique and maximal clique algorithms. Unlike many of those already developed in the literature, this new method aims to find convergence clubs endogenously without depending on a-priori classifications. In a Monte Carlo simulation study, the success of the method in finding convergence clubs is compared with a similar algorithm. Simulation results indicated that the proposed method perform better than the compared algorithm in most cases. In addition to the Monte Carlo, a new empirical evidence on the existence of convergence clubs is presented in the context of real data applications.

Keywords: Growth Economics, Convergence Hypothesis, Convergence Clubs, Maximum Clique Algorithm, Maximal Clique Algorithm.

JEL Classification: C32, O47.

*This work has been produced as a part of the research project (project no: 113K757) supported by The Scientific and Technological Research Council of Turkey (TUBITAK). Ege Yazgan and Thanasis Stengos would like to acknowledge financial support from TUBITAK, while Stengos would also like to thank SSHRC of Canada.

[†]Istanbul Bilgi University, email: fuadcan@gmail.com

[‡]University of Guelph, email: tstengos@uoguelph.ca

[§]Istanbul Bilgi University, email: ege.yazgan@bilgi.edu.tr

1 Introduction

One of the main predictions of (neoclassical) economic growth theory is that in the long run, all countries with similar technological characteristics would converge to a balanced growth path (steady state) equilibrium that will be entirely determined by the (exogenously) given growth rate of technical progress, which in turn would equal labour productivity growth. Hence economies with the same productivity would grow at the same rate and converge to the same equilibrium. This is the so called growth convergence hypothesis, which has been one of the main focal points of the empirical economic growth literature. In that context, a time series interpretation of the convergence hypothesis considers income gaps (or labour productivity gaps) between countries over time and analyzes whether these gaps would diminish, hence signifying convergence to a single steady state (equilibrium). On the other hand, if there are constant or increasing returns to capital, there may be a multiplicity of steady states (or absence of stable steady states) and a country's initial conditions will determine to which of these it will converge. In essence, convergence to a single steady state implies that however poor, a country will inevitably converge in the long run to a (prosperous) equilibrium shared by all. In the absence of such a single steady state, poor countries may only converge to a common equilibrium with other poor countries and will never catch up with the prosperous ones. The current debate on growth convergence as it has evolved over the last three decades, has been one of the most active research areas in economics and has taken the central role in the empirical growth literature. The main developments in the literature as they have evolved over time, mainly since the mid eighties and are summarized and presented in the Durlauf et al. (2005) survey. The earlier literature was based on the analysis of standard cross section/panel data and its main contribution was to identify all the main issues that have arisen in that context such as endogeneity, heterogeneity and nonlinearity. However, lately the emphasis has been on utilizing the existing data sets of long time series of GDP data compiled for most countries after WWII (and for a few developed countries going back to the nineteenth century). The resulting time series approach has built on the work of Bernard and Durlauf (1995, 1996) who have introduced time series interpretations of the convergence hypothesis that can be cast in terms of unit root and cointegration analysis.

In that strand of recent literature, Pesaran (2007) has extended the time series convergence concepts to the case where there is no requirement that the converging economies to be identical in all aspects including initial endowments. The main result is that for two economies to be convergent it is necessary that their output gap is stationary with a constant mean, irrespective of whether the individual country's output is trend stationary and/or contains unit root. Furthermore, testing for convergence in that case does not rely on using a benchmark country in order to define the output gaps that are used in the analysis and uses a pair-wise approach to test convergence. The issue of relying on a benchmark, also renders the analysis problematic as perceived leaders used as benchmark economies may not retain the leader title over the whole period of analysis. In that respect, Pesaran's (2007) pair-wise analysis becomes relevant. This analysis only considers the binary process of convergence (or lack of it) for all pairs out of a set of countries included in the initial group. The choice of this initial group is arbitrary and usually accomplished based on the data availability, geographic or economic developmental status. Therefore, the analysis has nothing to say how if one can also examine the issue of convergence to a common cluster that can be selected out of the initial group. Pesaran stated that "in principle, the convergence results from the analysis of pair-wise output gaps can be used to form "convergence clubs", but special care must be taken in addressing the specification search bias that such a strategy would entail" (Pesaran, 2007, p. 314). In other words the analysis so far, has mainly analyzed the issue of convergence between "country-pairs", but is mainly silent on how to proceed to classify countries as belonging to a common "country club".

Convergence to multiple steady states and the emergence of "convergence clubs" was put forward by various researchers in the literature, see Baumol (1986), Durlauf and Johnson (1995) and Galor (1996) to name a few. From a theoretical point of view, lack of convergence arises if there are constant or increasing returns to capital. In that case there may be a multiplicity of steady states (or absence of stable steady states) and a country's initial conditions will determine to which of these it will converge, see Azariadis and Drazen (1990) and Azariadis (1996). In essence, convergence to a single steady state implies that however poor, a country will inevitably converge to prosperity in the long run. In the absence of such a single steady state, poor countries may only converge to a common equilibrium with other poor countries and will never catch up with the prosperous ones. An alternative more recent and complete theoretical explanation for the

emergence of country clubs can be found in the Unified Growth Theory, see Galor and Weil (2000), Galor and Moav (2002) and Galor (2011). In that context the existence of multiple growth regimes arises naturally over time as economies differ in their respective phase of economic development. Hence, the differential timing of take-offs from stagnation to growth has segmented economies into three fundamental growth regimes: slowly growing economies in the vicinity of a Malthusian steady state, fast growing countries in a sustained-growth regime, and a third group of economies in transition from one regime to the other. However, the presence of multiple convergence clubs may be only a temporary phenomenon as endogenous forces may ultimately permit members of the Malthusian club to join the members of the sustained-growth club¹. From an empirical point of view, the broad evidence suggests that the process of convergence is not smooth but rather characterized by "start and stop" behavior. As argued by Johnson and Papageorgiou (2017) who have produced the most comprehensive survey on the subject of convergence up to date several mechanisms of divergence and convergence may be concurrently at work across countries in different stages of their development process.

In this paper we examine convergence to multiple steady states and the emergence of "convergence clubs" by introducing a new method that combines unit root testing within a $I(1)/I(0)$ framework with maximum and maximal clique approaches from the computer science graph theory to establish a set of statistical criteria for cluster formation². We will also offer an evaluation of the performance of our proposed method vis-a-vis other existing methods in the literature by means of a Monte Carlo simulation. To the best of our knowledge, this is the first time that the properties of such methods have been explored and analyzed in the literature. The paper is organized as follows. In the next section we will discuss the relevant literature on club formation. We will then proceed to discuss in detail the competing approaches that we will be investigating in section 3 and then we will present the description of the Monte Carlo design and discuss the data generating processes, evaluation procedures and Monte Carlo results in Section 4. In Section

¹In a recent paper Beylunioğlu et al. (2018) have analysed the evolution of output gaps between pairs of countries within a long memory framework using a Markov Switching Model that traces the evolution of the long memory parameter. It is then possible that several mechanisms of divergence and convergence are concurrently at work across countries in different stages of their development process and plotting the evolution of the long memory parameter would allow us to directly observe such behavior. Combining the above approach with the clustering formation mechanism that we examine in this paper could potentially allow one to study the evolution of club formation over time. Hence one could distinguish whether these clubs are permanent in nature (as in the case of a trap) or they are changing as the unified growth theory would suggest.

²The difference between maximum clique and maximal clique algorithms will be discussed in subsequent sections.

5 we will present the empirical results of the illustration of the method on real growth data and finally we will conclude.

2 Literature Review

The definition of a convergence club and the principle of clustering behind its formation gave rise to different empirical strategies to test the convergence hypothesis. However, the existing early methods were generally focused on the convergence of various *a-priori* defined homogeneous country groups which were assumed to share the same initial conditions. Baumol (1986) for example grouped countries with respect to political regimes (OECD membership, command economies and middle income countries), Chatterji (1992) allowed for clustering that based on initial income per capita levels and tested convergence cross-sectionally, while Durlauf and Johnson (1995) grouped countries using a regression tree method based on different variables such as initial income levels and literacy rates that determined the different "nodes" of the regression tree. An alternative approach to the cross-sectional notion of β -convergence in the context of cross-sectional was introduced by Bernard and Durlauf (1995, 1996) based on a time series framework that makes use of unit root and cointegration analysis, see Durlauf et al. (2005) for a comprehensive literature review for convergence hypothesis. Hausmann et al. (2005), similar to previous studies, by considering a priori grouping criteria such as initial incomes, found some evidence on convergence clubs by using time series methods.

In a time series context, Pesaran (2007) proposed a testing procedure that applies unit root tests to pairwise differences of the income per capita time series. This method relies on the use of unit root tests to all possible pairwise differences of the per capita income series in any given group of countries. Pesaran also considered different initial set of countries based on geographic characteristics for his pairwise method, but found no evidence on convergence clubs.

Similar to Durlauf and Johnson (1995), Hobijn and Franses (2000) (henceforth HF) proposed a panel data based approach for testing convergence. Contrary to the early attempts that relied on a two stage method that first assigns membership to a group and then considers whether this assignment is satisfied by the data, HF classifies countries into clusters of countries if they satisfy some criterion (desired convergence property). They clustered countries into subgroups by apply-

ing multivariate stationarity tests to panels consisting of pairwise differences of income per capita series and in contrast to Durlauf and Johnson (1995) they detected a larger number of small clubs. A different approach was proposed by Kapetanios (2003, 2008) who developed a method that is designed to endogenously classify stationary and nonstationary series by sequentially reducing the size of the null by removing series with the most evidence against the unit root null, classifying these series as stationary. The stopping point is when the unit root null is not rejected, such that all the remaining regions are classified as non-converging.

Using the HF methodology, Corrado et al. (2005) extended this method by allowing subgroups to vary over time and applied it to European regional sectoral data of agriculture, manufacturing and services. Corrado and Weeks (2011) extended the sequential HF approach to account for short time panels by using a bootstrapping modification and applied their method to study regional European convergence. The main contribution of HF is that it does not require an a-priori classification of country groups and detects group formation in an endogenous manner. A similar approach is advanced using the notion of σ -convergence by Phillips and Sul (2007) who developed an algorithm based on a log- t regression approach that clusters countries with a common unobserved factor in their variance. In the convergence literature, σ -convergence as opposed to β -convergence deals with the reduction in the variance of the cross country income distribution over time, see Quah (1996).

Following Pesaran (2007) and his testing procedure that applies unit root tests to pairwise differences of the income per capita time series convergence is reached when the proportion of rejections obtained from the pairwise unit root tests is greater than a certain threshold. He applied this method to country groups belonging to different geographical regions and found no evidence of convergence clubs. However, as is in most of the earlier studies, the country groups under consideration were defined subjectively a-priori without an endogenous clustering method. The current paper aims at developing a convergence analysis technique of cluster (club) formation that relies on pairwise testing both in the simpler $I(0)$ or $I(1)$ framework as in Pesaran (2007) combined with the maximum clique algorithm widely used in graph theory from the computer science literature, see Bron and Kerbosch (1973) and Konc and Janezic (2007). Rather than testing a-priori grouped country clusters, the method explores all convergent groups in a list of N countries that was previously subjected to pairwise convergence tests within a $I(0)/I(1)$ or a long

memory framework. Within a long memory framework this method has been introduced recently by Özkan et al. (2014) on a limited scale to study club formation among small (exogenously) defined groups of homogeneous countries. We propose to use this approach as a new endogenous cluster formation method for the all available countries and analyze and compare its properties with the existing endogenous cluster formation mechanisms of HF as they both rely on testing the time series properties of the mean function of output gaps as opposed to the variance (σ -convergence of Phillips and Sul (2007)).³ In our paper we will compare these two approaches, by means of an extensive Monte Carlo simulation study using evaluation criteria from the forecasting literature. This will be the first time that the properties of such mechanisms will be investigated and compared.

3 Methodology

The simple pairwise method and HF are both seeking convergence by searching *similarities* in movements of outcomes in the process of time. To this end both methods expect all pairs in a club to move around zero or a constant, in particular stationarity in difference of pairs. However there are several differences in approaches as well as the treatments of pairs. First, HF constructs clubs endogenously via a clustering algorithm that runs recursive stationarity tests. On the contrary, the pairwise method does not construct clubs, but tests the lists of clubs that are given exogenously. However our approach will combine pairwise testing with the maximum and maximal clique algorithms from computer science graph theory introduced by Özkan et al. (2014). As it will be argued below, there is a crucial theme in the construction of a single club, HF is a bottom up method that forms the clubs by adding countries one by one while the maximum clique method, by definition of employing the pairwise method, is a top down method that finds all the set of countries satisfying the definition of a club.

We proceed to present our proposed new convergence analysis technique that consists of pairwise testing as in Pesaran (2007) combined with a maximum clique algorithm widely used in graph theory from computer science literature. We first present the pairwise testing method and then

³In the last few years, there are a number of papers that have looked at (club) cluster formation in different research areas. Abbott and De Vita (2013), Fritsche and Kuzin (2011), Abbott et al. (2012), Kim and Rous (2012), Apergis and Padhi (2013), Yilmazkuday (2013) and Ikeno (2014) to name a few.

the procedure to find convergence clubs via the maximum clique technique. We will proceed to compare our proposed method with that of HF by means of an extensive Monte Carlo simulation study.

3.1 Pairwise Convergence Test

Suppose that the log GDP per capita series of country i and j at time t are as follows

$$Z_t^{ij} = y_t^i - y_t^j = \beta + \varepsilon_t \sim I(d), \quad i = 1, \dots, N-1, \quad j = i+1, \dots, N, \quad t = 1, \dots, T$$

where T is the length of time interval, N is the number of countries and y_t^i and y_t^j denotes the log GDP per capita series of i and j . ε_t stands for the disturbance term and $d \in \{0, 1\}$ represents the integration of the series. Here β can represent a constant or a function of time as well. (see Stengos and Yazgan (2014)). Since the difference series are either stationary or non-stationary, that will determine if the pair is convergent or not. For instance if $\varepsilon_t \sim I(0)$, the two log GDP per capita series will be drifting together overtime and in that case it is appropriate to assert that countries i and j are convergent. On the other hand, if $\varepsilon_t \sim I(1)$, a nonstationary process would indicate that the log difference series between i and j is nonstationary and the two log GDP per capita series would be drifting apart over time, indicating that countries i and j are not converging.

Determining convergence by applying unit root tests on differences between GDP per capita series characterizes the time series based approach on convergence applied in many different contexts since Bernard and Durlauf (1995, 1996), see Durlauf et al (2005) for a comprehensive survey. However, when there are more than two countries, there is still uncertainty in determining whether the countries are converging altogether to a steady state. In the literature, the main approach centers on testing if all countries in the group are converging to the group average or a chosen country as a benchmark (generally United States), hence applying unit root tests to the pairwise differences of each group member with the average or the selected benchmark country. Alternatively, another approach is to apply multivariate stationarity tests to determine convergence. The former approach is criticized for the arbitrariness in choosing the benchmark country or the country average, while the latter is not preferred because of the difficulties in applying it to large

groups.

The pairwise method developed by Pesaran (2007) can offer a remedy to both of the above difficulties. According to this approach, if one tests for convergence of a group of N countries, all $N(N - 1)/2$ pairs are subjected to unit root testing. Pesaran (2007) showed that, if a group of N countries are non-convergent, the rejection rate of the null hypothesis of non-stationarity ($H_0 : Z_t \sim I(1)$) calculated by $N(N - 1)/2$ tests is equal to the nominal size of the individual tests, i.e. the probability of Type 1 error. More specifically, it is shown that under the null hypothesis of N countries being non-convergent, the rejection rate of individual tests converges to the nominal size, α , as N and $T \rightarrow \infty$, even though individual tests are not independent cross-sectionally. Since the null hypothesis in this case is non-convergence (divergence) of N countries, in order to find evidence in favor of the null, it is enough to show that the proportion of rejections over $N(N - 1)/2$ tests is not larger than the significance level of individual tests. In that case for example, if the significance level is 5%, the proportion of rejections must not exceed 0.05⁴. To summarize, rejection rates of $H_0 : Z_t \sim I(1)$, higher than a given significance level in a given application would imply evidence against the non-convergence (divergence) null hypothesis in favor of the convergence alternative. On the other hand, rejection rates lower or close to the employed significance level will provide evidence for the non-rejection (validity) of divergence.

3.2 Maximum Clique Method for Finding Convergence Clubs

The maximum clique method that we present in this subsection, combines the maximum clique algorithm of graph theory with the previously described pairwise convergence tests of $H_0 : Z_t \sim I(1)$. Rather than testing a priori grouped country clusters, the method explores all convergent groups in a list of N countries that was previously subjected to pairwise convergence tests. In this sense, the method is an endogenous extension of Pesaran (2007) similar to the one proposed by HF.

The method consists of two steps. First, all possible pairwise differences of N countries are subjected to unit root tests where the null denotes a unit root process as evidence of non-convergence.

If the rejection rate obtained from $N(N - 1)/2$ tests is well above the significance level, that would

⁴No doubt, nominal size of the tests may differ from the significance level. In applications, the power of the tests used relative to size distortions should be given attention. Another matter to be attentive to is the fact that the rejection rate would converge to α in the limit and that of course would not be the case if N and T are relatively small in a given application.

be evidence against the null hypothesis of non-convergence (divergence) hypothesis in favor of the alternative of convergence and the list of N countries will be taken to form a convergent group. If this club involves all examined countries, then all countries are said to be convergent and we do not go any further in seeking out the presence of convergence clubs. However, as shown in Pesaran (2007), Dufrénot et al. (2012) and Stengos and Yazgan (2014) it is very unlikely that, by examining all countries as a single group, one will find evidence of convergence for all with pairwise testing. Nevertheless, if a subgroup of countries is found convergent via pairwise method, then it can be said that this subgroup constitutes a convergence club. The main challenge, as indicated above, is to find a method to determine this subgroup rather than relying on a-priori classifications. In the second step we undertake this task.

Assume that \mathcal{U} denotes the set of all countries. Hence, by definition, the cardinality of \mathcal{U} is equal to N ; mathematically if $\#(\cdot)$ denotes the cardinality, we have $\#(\mathcal{U}) = N$. Moreover, suppose that \mathcal{E} is a subset of \mathcal{U} . In this case, in order \mathcal{E} to be a convergence club, all binary combinations obtained from the elements of \mathcal{E} should satisfy the stationarity hypothesis in the pairwise tests. In other words the null of $Z_t \sim I(1)$ should be rejected for all $m(m-1)/2$ pairs, where $\#(\mathcal{E}) = m < N$.

In the second step, from the $N(N-1)/2$ test results, the objective is to find a class of subsets \mathcal{G} for which all subsets, e.g. \mathcal{E} , satisfy pairwise convergence property. Mathematically, let \mathcal{G} denotes the class of all subsets satisfying the desired pairwise (stationarity) property. Then the problem is

$$\mathcal{G} := \{\mathcal{E} : \forall i, j \in \mathcal{E}, t(Z^{ij}) = 1\}$$

where $Z^{ij} = y^i - y^j$, $t(\cdot)$ is the test result of the series in the bracelet and takes the value of 1 for a convergent pair, i, j and 0 otherwise.⁵ Hence, the problem can be expressed as

⁵Notice that in order to satisfy the property explained above, \mathcal{E} to be a convergence group, all pairs $i, j \in \mathcal{E}$, $i \neq j$ should satisfy the convergence property. Conversely, we require the non-rejection rate of $H_0 : Z_t \sim I(1)$, which denotes divergence, to be zero, which is a much more stringent condition since it does not allow Type 1 error. Similarly, we expect the rejection of $H_0 : Z_t \sim I(1)$ to be unity if convergence holds. One can relax this condition by allowing rejection rate up to a given level. As explained above Pesaran (2007) shows that when the number countries in a club, $N \rightarrow \infty$ the pairwise rejection rates approaches to significance level of stationarity tests. Although adopting Pesaran's approach is possible for large number of countries, it is questionable for the club sizes considered in the simulations. It is also worth noting that allowing the rejection rate up to a given level (say at the nominal size of a significance test) may have an impact on how much Type 1 error created by the unit root test would carry over in club formation. However, since our main goal is to compare different club formations mechanisms with Monte Carlo simulations relying on the same unit root tests, all these methods under comparison will be on the same footing.

$$\arg \max_{\mathcal{G}} \{ \#(\mathcal{E}) : \mathcal{E} \in \mathcal{G} \}.$$

In graph theory terms, countries become vertices, the test result (rejecting or not rejecting pairs) of a pair become edges, and as such the set of all vertices and edges constitutes an undirected graph. If an undirected graph has edges between all vertices then the graph is said to be complete. If there is a subset of an undirected graph having all properties of a complete graph, the subset is so called a *clique*. Therefore, in our case, all convergence clubs of a country list can be expressed as cliques. Solving the problem defined above is known as finding maximum cliques.

Pairwise test results form an undirected graph and accordingly, countries and test results determine the vertices and edges respectively. Hence, the problem becomes to find a subgraph with the maximum number of vertices, where an edge is defined between two vertices, or in other words, a maximum clique. Figure 1 and 2 presents the notions mentioned above.

Finding a maximum clique can be too hard from a computational point of view. The computational complexity of solution to maximum clique problem is known as NP-Complete whose brute-force solution requires $2^N - \binom{N}{2} - N - 1$ trials. First, Bron and Kerbosch (1973) developed an algorithm to solve the problem in exponential time. In the recent literature, various planar graph algorithms have been developed that enables the problem to be solved in polynomial time. In this study, we will employ the branch and bound algorithm proposed by Konc and Janezic (2007) which is an improved branch-and-bound algorithm that ends in polynomial time. We will apply this method using the *R* programming language with the *igraph* package.

We should note that, the maximum clique method is not a conclusive technique. In other words, it does not cluster the country list into subgroups, but finds club(s) having a maximum number of elements. Hence we offer the following clustering algorithm to detect convergence clubs.

1. Apply the desired stationarity test to all Z^{ij} such that $i, j \in \mathcal{U}$ and $i \neq j$.
2. Test the unit root null hypothesis. The resulting variable takes the value of 0 if null of non-stationarity (unit root) is not rejected, 1 if it is rejected (evidence for stationarity).
3. Construct adjacency matrix from the resulting variable values obtained in (2).

4. Find maximum clique(s) from the adjacency matrix via the algorithm proposed by Konc and Janezic (2007).
5. The group of countries in the clique is labeled as a convergence club. Eliminate respective rows and columns of the countries from the adjacency matrix. And step back to (4). Stop if all the rows and columns are eliminated from adjacency matrix.

3.3 The HF Method

As mentioned in the introduction, the method presented above bears certain similarities to the endogenous cluster analysis proposed by HF. Hence, it is important to compare the accuracy of our method with HF. To this end, we will first present HF and review both method by means of a simulation comparison in the following two subsections⁶. An important difference between the maximum clique method described above and the HF approach is that the former is based on testing the null of a unit root (divergence), whereas in the latter the null hypothesis is stationarity (convergence).

HF is a clustering algorithm that applies multivariate KPSS tests recursively to panels enlarged by a series in each iteration. More generally, the algorithm allows a new country to enter the convergence group until null hypothesis of stationarity is rejected. HF relies on two definitions of convergence clubs. The first of these, perfect convergence, requires club members to have statistically equal GDP per capita series. Perfect convergence occurs if the pairwise difference of the club members' output series are stationary around a zero mean. This definition of convergence indicates a more stringent state, since it ignores catching-up possibilities or other differences stemming from initial conditions. The second definition of convergence, the so called relative convergence, describes similar movements in the output series over time irrespective of initial conditions, i.e. the pairwise differences follow the non-zero mean stationarity property.

As mentioned above, HF determines convergence of a group by utilizing multivariate KPSS tests. The method is based on the construction of a panel containing pairwise differences of

⁶As remarked in the literature review, other than HF, another method developed by Phillips and Sul (2007) stands out by means of not requiring a priori classification of countries. However, we exclude this method for the following reason. Unlike HF and our proposed pairwise maximum clique method, Phillips and Sul (2007) is based on σ type convergence. The method depends on the definition of convergence by means of reduction of variance over time and thus convergence of series to a steady state. Therefore, it is not appropriate to compare this method with HF and the method developed in this study as both of the latter deal with convergence of the mean (function) of the series.

consecutive series, then it applies KPSS test to this panel. In this manner, to test if country group $C = \{c_{n_1}, c_{n_2}, \dots, c_{n_p} : n_p < N\}$ is converging, the panel $\mathbf{x}_t^C \equiv \mathbf{M}_p \mathbf{y}_t^C$ is defined where \mathbf{M}_p and $\mathbf{y}_t^C \in \mathbb{R}^p$ are as follows.

$$\mathbf{M}_p = \begin{bmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & -1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{y}_t^C = \begin{bmatrix} y_{1t} \\ \vdots \\ \vdots \\ y_{kt} \end{bmatrix}.$$

Here, \mathbf{x}_t^C denotes the matrix of consecutive differences of incomes, $Z_t^{(i-1)i}$, $\forall i \in \{n_1, \dots, n_p\}$, and the stationarity test applied to \mathbf{x}_t^C determines whether or not country group C constitutes a convergence club. HF tests perfect and relative convergence separately by employing two respective multivariate KPSS tests. Convergence clubs from N countries are clustered via the following algorithm.

HF Algorithm:

1. Initially, set each countries as a club by defining $l_i = \{i\}$ for $i = 1, \dots, N$.
2. For each $i < j$, construct \mathbf{y}_t^C where $C = l_i \cup l_j$. Through these matrices, apply the multivariate KPSS test to \mathbf{x}_t^C . If null hypothesis of stationarity is rejected for all i, j , reject convergence hypothesis and stop. If it is not rejected for any pair, i, j proceed to next step.
3. Choose i, j that is tested to have largest p-value from the KPSS test in the previous step. For $i < j$, redefine l_i as $l_i = l_i \cup l_j$ and set $l_j = \emptyset$. Step back to (2).
4. Label non-empty sets obtained with more than one member as convergence clubs.

3.4 Comparison of Methods

HF is a method that relies on a "bottom up" algorithm that clusters groups one by one. On the contrary, the maximum clique method relies on a "top-down" process that detects all subsets satisfying club properties. Other than clustering, there is a substantial difference in testing conver-

gence. To determine whether a set of countries is convergent, HF applies multivariate stationarity test to panels comprised of consecutive pairwise difference series set elements and confirms convergence if the null hypothesis of stationarity of the panel is not rejected. However, the panels do not include all possible pairwise differences but only differences of consecutive pairs. For example, if we want to test the convergence of countries 1,2,3 and 7, a panel consisting of Z^{12} , Z^{23} and Z^{37} is subjected to the test, and if stationarity cannot be rejected the panel is then augmented with series other than 1, 2, 3 and 7, each added separately. If then for each of these additional panels the stationarity null is rejected, then these four countries are said to be convergent. On the other hand, our proposed pairwise method depends on a different definition of clubs, so that for m countries to be convergent, we need to achieve rejection of the null of a unit root for all $m(m - 1)/2$ pairs. Hence, in order for the list of countries in the previous example to form a convergence club, the rejection rate of $4(4 - 1)/2 = 6$ pairs from unit root tests should exceed some significance level.

3.5 Monte Carlo Structure

In this subsection, we will discuss the data generating processes that is used in our Monte Carlo study. We generated various types of data to conduct the evaluation of the clustering methods that we compare in order to determine factors and sources leading to success and failure. The data sets are classified in two groups. In the first group we include single club and many non-convergent pairs, while the ones in the second group include multiple clubs together with only some non-convergent pairs. In the following parts of this subsection, we will present the data generating processes and evaluation procedures employed in this study.

3.5.1 Data Generating Processes

The simulation assumes that the log GDP series is given as follows.

$$y_{it} = c_i + \gamma_i f_t + \epsilon_{it} \tag{1}$$

where $\epsilon_{it} \sim I(0)$ is the error term and f_t is the common factor which affects all countries the same way (such as technology). If we assume non-stationarity of the factor, a pair of countries

can only be convergent if both countries utilize the factor likewise. This can be possible if the country specific constants, γ_i that measure that effect are equal. In other words, for the pair i and j , if $\gamma_i = \gamma_j$, f_t is canceled out and $y_{it} - y_{jt}$ becomes $c_i - c_j + \epsilon_{it} - \epsilon_{jt}$. In this case, since the error terms are assumed to be stationary, we have $c_i - c_j + \epsilon_{it} - \epsilon_{jt} \sim I(0)$ and the pair i and j would be convergent by definition. Likewise, for any subset of countries having equal γ_i terms, all pairwise difference series in that subset would be stationary and hence these countries would constitute a convergence club. On the other hand, the constants, c_i , are country specific and are generated once for all data sets.

The non-stationarity of f_t is modeled under an ARIMA process following below

$$f_t = f_{t-1} + v_t, \quad v_t = \rho_v v_{t-1} + e_t, \quad e_t \sim iid N(0, 1 - \rho_v^2)$$

where we allow $\rho_v = \{0.2, 0.6\}$ as separate cases. Besides, we also allow the error term of the log GDP series in equation (1) to have serial dependence, following the specification below,

$$\epsilon_{it} = \rho_i \epsilon_{i,t-1} + v_{it}, \quad v_{it} \sim iid N(0, \sigma_{v_i}^2 (1 - \rho_i^2)).$$

Above, the error terms, v_{it} are i.i.d. distributed Normal random variables. Here, the autoregressive coefficient ρ_i and $\sigma_{v_i}^2$ are country specific and invariant among the data sets. To be more precise, before proceeding to data generation, we generated the coefficients to have the following property.

$$\sigma_{v_i}^2 \sim iid U[0.5, 1.5], \quad \rho_i \sim iid U[0.2, 0.6]$$

To generate a single club containing data set, the coefficients of m convergent countries, are assumed to be $\gamma_i = \gamma_j = 1$. For the remaining $(N - m)$ countries, γ_i generated randomly as $\gamma_i \sim iid \mathcal{X}_{\kappa_i}^2$. It is worth noting that γ_i are generated once, yet, when the number of club members (m) is 10 instead of 5 for example, arbitrarily selected 5 of the remaining coefficients are substituted with 1 to allow them to be convergent. Lastly, we also generate country specific constants as $c_i \sim iid \mathcal{X}_{\kappa_i}^2$.

For multiple clubs, the club sizes, m 's associated with each club, when the number of clubs (k) and the number of countries (N) are given, are determined by allowing different varieties in

club sizes and in a manner in which 2 single non-convergent countries that do not belong to any club are present in the data. Additionally, for a given k and N we selected, the clubs sizes m 's are randomly drawn from a Poisson distribution with a rate of $\lambda = N/k$. For each N , random draws are repeated k times⁷.

For both procedures, γ_i 's are equal for countries constituting a club, but unequal among clubs. In particular, γ_i are chosen from $\{1, 2, 4, 5, 6, 9, 11, 12, 13, 14, 15, 16, 18, 21, 22, 24, 25, 26, 28, 30\}$ and the remaining coefficients of non-convergent countries are generated as in single clubs. The simulations are repeated 10000 times for various time intervals; number of countries, clubs and club members. In particular, simulations are performed using different combinations of $T = \{50, 75, 100, 200\}$ time intervals, $N = \{10, 20, 30, 40\}$ count of countries and $k = \{1, 2, 3, 4, 5, 6, 7, 8\}$ number of clubs.

3.5.2 Testing and Evaluating Procedures

We start by generating the repeated data sets based on the number of replications following the above specifications and choice of parameters. Both methods are applied to each generated data sets and the resulting club(s) obtained via both methods are evaluated by comparing the predicted club formation of each method with the actual club formation from the data generating processes described above. The general evaluation of the success of each method is considered for each data type.

To the best of our knowledge there is no other comparable Monte Carlo study in the literature that evaluates clustering methods in the same context as we do here. We will make use of some statistics from other fields to evaluate the maximum clique algorithm and HF and the methods that we will propose differ for single club to the multiple club cases. We will first present the single club evaluation methods, for which we will utilize two different evaluation tools. The first one is the Kupiers Score (KS), while the second one is the Pesaran and Timmermann (1992) Test Statistics (PT) commonly used in the forecasting times series literature for the evaluation of sign forecasts. It is worth noting that sign forecasts are used for predicting whether an underlying series would appreciate (increase) or depreciate (decrease) relative to a benchmark. In our case, success in detecting a country's membership in a club is equivalent to success in forecasting the

⁷Obviously we did not allow the sum of m to exceed N , if this happens we redrew the last club size.

sign of a time series.

Since success of bidirectional results such as upside and downside movement or membership in a club, can occur randomly⁸, KS takes true forecasts and false alarms into account separately. For instance, if we are evaluating a forecast of a bad event or calamity in economics, an estimate of false alarms would help us avoid the issue of scare-mongering. Here, KS is defined as $H - F$ where

$$H = \frac{II}{II + IO}, \text{ and } F = \frac{OI}{OI + OO}$$

The capital letters "I" and "O" in the above formulae refers to whether the country under investigation is a member ("in" the club) or not ("out" of the club). Regarding the order of the letters, the first letter indicates whether the country is found to be a member in the experiment, while the second letter denotes its actual membership state (i.e. whether the country is actually in the club or not). Therefore, "II" indicates that the country, as a member of the club is correctly identified whereas, "OO" denotes that that the country, as an outsider of the club is also correctly identified. Furthermore, IO indicates that a country is detected to be in the club, while actually it is not (false detection). Similarly OI refers to the case where the country is misclassified as being outside, even though it is a member of the club (false alarm). The ratio H captures the rate of "correct hits" in detecting club membership, whereas F denotes the "false alarm" rate, that is the rate of false exclusions.

As in the case of sign prediction in the forecasting literature, success can be the outcome of a pure chance probability event of 0.5. Hence, to test the statistical significance of KS, we will employ the following PT statistic,

$$PT = \frac{\hat{P} - \hat{P}^*}{[\hat{V}(\hat{P}) - \hat{V}(\hat{P}^*)]^{\frac{1}{2}}} \sim N(0, 1),$$

where \hat{P} refers to the proportion of correct predictions (correct detections of countries as being a member or non member) over all predictions (N countries), and \hat{P}^* denotes the proportion of correct detections under the hypothesis that the detections and actual occurrences are independent.(where success is a random event of probability 0.5), while $\hat{V}(\hat{P})$ and $\hat{V}(\hat{P}^*)$ stand for the

⁸This is similar to expecting an unbiased coin to come up heads with 50% probability.

variances of \widehat{P} and \widehat{P}^* respectively.

In simulations involving multiple clubs, it is not possible to use KS and the PT statistic due to the more complicated nature of the success/failure classification which is no longer binary as in the case of the single club case. In the situation of a single club, a country can be either detected (correctly or incorrectly) to be a member of this single club or not. In the multiple club case, on the other hand there are more than two distinct cases for the actual membership state: the country can be either a member of the correct club, belong to the "wrong" club, or not be a member of any club.

To confront this problem, in the case of multiple clubs we will use a much stricter criterion by counting the successful cases in our simulations in which *all* countries are detected correctly to be in their correct positions. In other words, we do not evaluate success as a binary outcome, country by country as in the case of a single club in each replication. Instead, we will look at the overall results in each replication. If, in one replication, all countries are placed correctly in their correct position we will consider this as one successful outcome out of a total of 10000 replications and as fail otherwise.

3.6 Simulation Results

We now proceed to discuss the findings of the simulations based on the data generating processes of club formation discussed above. We will first discuss the comparison between the pair-wise unit root based approach augmented with the maximum clique algorithm and the HF approach based on multivariate KPSS testing for the single club case and then the multiple club case.

3.6.1 Single Club Results

The results are presented in Tables 1 and 2. Table 1 presents two choices of the number of club members ($m = \{3, 5\}$), whereas Table 2 choices for ($m = \{7, 10\}$). In each table there are four choices of the total number of countries involved N , ($N = 10, 20, 30, 40$); four choices of time span ($T = 50, 75, 100, 200$) for the analysis that would mimic the real data time span availability; and two choices of the persistence parameter ($\rho_v = \{0.2, 0.6\}$). The pair-wise unit root testing approach is combined with the maximum clique algorithm using an ADF pair-wise test⁹, while HF

⁹Even though we initially considered 3 different ways that the pair-wise unit root testing approach could be combined with the maximum clique algorithm, one that uses an ADF test, the second using the ADF with a GLS

is based on the multivariate KPSS testing procedure. Among the three versions of the pair-wise approach, the ADF one gives better and more consistent results that overall outperform its all competitors including HF. The last set of columns for example in Table 2, for the $H - F$ results (the "correct hit" ratio net of "false alarms", indicates that with $m = 10, N = 40, T = 100$ and $\rho = 0.6$, ADF with a 0.951 KS outperforms the others including the HF method that has a KS of 0.845. Similarly the the PT test statistics yield 579.5 for the pair-wise ADF test versus 552.2 for HF. Note, that the rejections of the null hypothesis of random success outcomes are higher with the PT test for both methods (slightly more so for the pair-wise ADF approach). Our method conducting the pair-wise analysis based on ADF alone combined with the maximum clique algorithm will be henceforth denoted as MCL.

3.6.2 Multiple Club Results

The results for the multiple club case are presented in Table 3 and 4. The multiple clubs cases involve classifications with $k = 2, 3, 4, 5, 6, 7, 8$ and $N = 10, 20, 30, 40$ for $T = 50, 75, 100, 200$. In Table 3 the club sizes (m 's) for multiple clubs are chosen in several combination varieties, whereas in Table 4 they are randomly picked from a Poisson distribution with a rate of $\lambda = N/k$. The m 's associated with each club are explicitly listed in the third columns of Table 3 and 4 for each k and N . We only present the ADF version of the pair-wise method as it was clear from the previous single club analysis that the other two versions were outperformed by the simple ADF. Again, the pair-wise MCL algorithm outperforms HF in the majority of cases and especially when N (the pool of countries available) and T , the time span increases, both in the case of the presence of a constant or not. For example, with $N = 20, k = 5, T = 100$ and $\rho_v = 0.6$ the pair-wise MCL detects 89.9% correct classifications without the constant and 80.0% cases with the constant DGP, while HF detects 53.6% and 53.1% such cases respectively. The results suggest that in terms of accuracy the ADF-maximum clique augmented pairwise method does quite well in detecting correctly the presence of clubs or clusters of countries. This gives us confidence that using the above method to real data would provide us with useful insights about how countries over time collect themselves into different groups and club formations of similar characteristic as far as economic activity is concerned.

corrected unit root test and a third one using the KPSS test, to conserve space we only report the one that uses the ADF test as it outperforms the other alternatives.

4 Real Data Application: Growth Convergence

in this section to apply the club formation methods analyzed earlier using the GDP per capita data from the Penn World Tables (PWT) and the Maddison Project Databases.¹⁰

The data from Maddison Project go back to very early years for some countries and data availability increases mainly after 1930. In fact, after 1930, 36 countries have no missing data point between 1930 - 2010 and the largest number of countries available for the period of 1950-2010 is 95. On the other hand, PWT data are available for the period of 1950 - 2014 for 55 countries only. In addition to these three data groups, we also consider three different types of pre-determined country groups that are considered important in the empirical growth literature: Europe, the Group of Seven (G7) and the S&P Emerging Markets classification group. Table 5 displays the list of the countries covered under each data classification. In our applications given in Table 6, we used different combinations of these pre-determined data groups.

In the application we will follow a slightly different clustering algorithm, maximal clique, that is an extension of the one we used in the multiple club simulations. In the simulations we used an iterative methodology that finds and tags the maximum clique (the clique with maximum number of members) of a given graph as a club, excludes the members from the initial list of countries and applies the same procedure to the rest of the list iteratively. This very strict procedure fits the purpose of Monte Carlo simulations where the true club formation mechanism is known and one seeks to obtain perfect club detection. In that case, misclassifications would arise from statistical sampling errors due to the adopted testing procedure. However, with real data such a procedure may lead to club formations beyond the largest club that are incompletely characterized. For instance, if we go back to the example presented in figure 1 and 2, the maximum clique algorithm would only detect the largest clique (countries 1 to 7 with seven countries) among several others existing in the graph. Hence it would disregard smaller ones (e.g. the one with members 6, 7, 8 and 9) which may be meaningful in economic terms. In this case the above iterative procedure will choose and tag the larger clique (1 - 7) in the first iteration, breaking off members 6 and 7 from the smaller 4-member club. Figure 3 illustrates this case.

¹⁰The Maddison-Project: <http://www.ggd.net/maddison/maddison-project/home.htm>, 2013 version,; Penn World Tables: <https://www.rug.nl/ggd/productivity/pwt/>, version 9.0.

To overcome this problem we will make use of another notion from graph theory, *maximal clique*. A maximal clique can be defined as a clique that is not a subset of any other clique. Thus, detecting all maximal cliques in a group of N countries provides us the list of all convergence clubs excluding their subsets. In other words, the set of all convergence clubs, C is a subset of G which is not a subset of any other $E \in G$. Hence compared to the procedure we applied in the multiple club simulations, the maximal clique algorithm does not disregard smaller clubs, but lists them as potential convergence clubs¹¹. As illustrated in figure 4, countries 6 and 7 are detected to be in two different clubs and counted accordingly as members of these two clubs, since they may share important characteristics about their economies. Note that in our Monte Carlo setting where there were only non-overlapping clubs the maximum and maximal clique algorithms coincide.

Table 6 displays the number clubs that are found by HF and maximal-clique algorithms¹² using a 5 percent significance level, where for our MCL pairwise method the lags for the application of the ADF tests are selected automatically by the Akaike Information Criterion. Examining for example the results obtained for the 1930 group ($N = 36$), the MCL method finds 14 clubs with 2 countries (# 2), 21 clubs with 3 countries (# 3) and finally 7 clubs with 4 countries (# 4). For the HF method there are 5, 2 and 5 clubs with # 2, # 3 and # 4 members respectively. Recall that the search of these convergence clubs is done over the total of 315 ($= N(N - 1)/2 = 36(35)/2$) country pairs. As explained above, MCL does not exclude the possibility of overlapping countries in different clubs with the same number of countries. In that case for instance, at least some of the 7 clubs with 4 countries would be expected to include some of the same countries. On the other hand HF as a result of its algorithm categorizes the list of all countries as convergence clubs with distinct (non-overlapping) elements. Therefore, the 5 clubs with 4 countries contain necessarily distinct countries.

Figures 5 to 12 illustrates some of the club formations over different data sets. For the data combining Euro and S&P, MCL finds Morocco, Portugal, Spain and HF finds Austria, Egypt, Finland, Indonesia and Italy as constituting a club. For G7 + Europe MCL detects USA, Netherlands, Germany and USA, France, Germany as different alternatives, whereas HF finds Bulgaria, Germany, Netherlands, Poland, UK; and finally for G7 + S&P MCL finds USA, France, Ger-

¹¹To detect maximal cliques, we used `maximal.cliques` command in the `igraph` package

¹²Although maximal clique algorithm differs from the maximum clique algorithm used in the simulations, we will continue to use the same MCL acronym as they are both very closely related.

many and USA, France, Italy as alternatives, whereas HF finds Egypt, France, and Indonesia as constituting a convergence club. Overall, it seems that the MCL algorithm is allowing for more homogeneity in the clubs that are formed, since the overlapping groups that enter more than one club act as common factors that are shared by all. On the other hand the HF method emphasizes distinctness and as such some of the clubs that are formed are difficult to interpret from an economic point of view as they mainly pertain to economies with different structures.

5 Conclusions

In this paper we have introduced a new method that combines unit root testing within a $I(1)/I(0)$ framework with the maximum clique and maximal clique approaches of graph theory to establish a set of statistical criteria for cluster formation. We offer an evaluation of the performance of our proposed MCL method vis-a-vis the HF method in the literature, that is closer in spirit to our approach, by means of a Monte Carlo simulation. To the best of our knowledge, this is the first time that the properties of these methods have been explored and analyzed in the literature. In the application we encountered almost the same patterns as in the single club simulations. The results with real data reveal that the MCL method based on the maximal clique algorithm provides more meaningful club classifications than HF, even though both methods result in similar numbers of clubs being formed. However, it is worth noting that both MCL and HF are methods based on statistical testing and the accuracy of the club formation results will depend on the power properties of the underlying test statistics used.

We have compared our proposed MCL method that is a top down approach to the simple bottom up approach of HF. In future research we also plan to extend our analysis to also cover alternative variants of the HF approach as in Corrado and Weeks (2011) and the alternative method proposed by Kapetanios (2008). It is worth noting that the methods proposed and examined here are based on an analysis of the mean function and they do not account for σ -convergence as in Phillips and Sul (2007). Examining the properties of the latter method is also left for future research. Finally, it would be interesting to examine the evolution of club formation over time by combining the methodology that we have introduced here and the Markov Switching model within a long memory framework of Beylunioğlu et al. (2018) to study the evolution of club formation

over time and potentially distinguish whether these clubs are permanent in nature (as in the case of poverty traps) or changing between states as the unified growth theory would suggest.

6 Tables and Figures

Table 2: Single Club: Kupiers Scores and PT Statistics, $m = \{7, 10\}$, 5% significance level

m	N	ρ	T	No Constant								With Constant							
				H		F		KS		PT		H		F		KS		PT	
				MCL	HF	MCL	HF	MCL	HF	MCL	HF	MCL	HF	MCL	HF	MCL	HF	MCL	HF
7	20	0.2	50	0.991	0.837	0.025	0.015	0.967	0.821	428.8	381.4	0.849	0.653	0.055	0.199	0.794	0.454	359.4	202.0
			75	0.997	0.897	0.029	0.019	0.968	0.878	428.1	399.2	0.977	0.748	0.070	0.225	0.906	0.523	396.8	227.2
			100	0.997	0.927	0.035	0.021	0.962	0.905	424.6	408.3	0.984	0.804	0.081	0.239	0.904	0.565	394.0	243.0
		200	0.997	0.962	0.045	0.024	0.952	0.938	418.7	418.9	0.990	0.910	0.103	0.206	0.887	0.704	384.1	301.8	
		50	0.991	0.840	0.028	0.012	0.963	0.828	426.9	384.6	0.862	0.676	0.027	0.180	0.835	0.497	382.2	221.6	
		75	0.996	0.899	0.030	0.018	0.966	0.882	427.0	401.2	0.986	0.765	0.029	0.209	0.957	0.556	424.4	241.6	
	100	0.996	0.932	0.036	0.020	0.960	0.912	423.6	410.9	0.994	0.820	0.032	0.223	0.962	0.597	425.6	257.1		
	200	0.997	0.963	0.035	0.021	0.962	0.942	424.7	420.9	0.996	0.913	0.034	0.198	0.961	0.714	424.6	306.4		
	50	0.984	0.801	0.041	0.020	0.943	0.781	497.6	450.3	0.824	0.556	0.094	0.179	0.730	0.377	383.7	197.2		
	75	0.991	0.865	0.055	0.022	0.935	0.843	486.2	472.1	0.970	0.635	0.124	0.209	0.845	0.426	419.3	213.7		
	100	0.992	0.898	0.067	0.026	0.925	0.872	475.9	480.1	0.989	0.694	0.132	0.217	0.857	0.477	421.4	235.4		
	200	0.994	0.945	0.089	0.027	0.905	0.918	457.3	496.7	0.995	0.808	0.148	0.210	0.848	0.597	412.9	290.2		
	50	0.984	0.818	0.038	0.014	0.947	0.804	501.0	464.0	0.835	0.578	0.066	0.165	0.769	0.414	414.1	218.2		
	75	0.992	0.880	0.044	0.017	0.947	0.863	497.0	484.1	0.970	0.658	0.092	0.192	0.878	0.466	445.5	235.2		
	100	0.992	0.911	0.049	0.019	0.943	0.892	492.3	493.0	0.986	0.719	0.098	0.201	0.888	0.518	446.9	256.9		
	200	0.994	0.950	0.066	0.024	0.928	0.927	477.6	502.6	0.995	0.841	0.108	0.194	0.888	0.648	442.7	315.6		
	50	0.980	0.795	0.033	0.016	0.947	0.779	570.2	520.5	0.773	0.501	0.087	0.152	0.686	0.349	406.8	205.1		
	75	0.989	0.858	0.045	0.020	0.944	0.838	557.0	540.7	0.953	0.589	0.102	0.178	0.851	0.411	471.4	227.9		
	100	0.990	0.892	0.051	0.019	0.939	0.873	549.4	555.9	0.976	0.656	0.113	0.181	0.863	0.475	470.2	258.9		
	200	0.993	0.940	0.069	0.025	0.924	0.915	527.2	566.7	0.990	0.771	0.122	0.183	0.869	0.588	468.4	312.8		
	50	0.982	0.809	0.033	0.012	0.949	0.798	571.4	534.6	0.813	0.530	0.054	0.140	0.759	0.390	467.8	230.8		
	75	0.989	0.873	0.037	0.015	0.952	0.858	568.3	555.1	0.966	0.613	0.068	0.165	0.898	0.447	516.4	250.3		
	100	0.991	0.905	0.043	0.017	0.949	0.887	560.8	564.1	0.984	0.681	0.076	0.173	0.908	0.509	514.6	278.1		
	200	0.992	0.950	0.054	0.022	0.938	0.928	546.3	574.4	0.992	0.814	0.083	0.182	0.910	0.632	510.6	333.6		
10	20	0.2	50	0.990	0.739	0.024	0.015	0.966	0.725	431.9	334.4	0.824	0.598	0.056	0.189	0.768	0.409	346.0	187.3
			75	0.998	0.834	0.029	0.019	0.969	0.815	433.6	368.6	0.980	0.696	0.072	0.219	0.908	0.477	406.5	214.1
			100	0.998	0.878	0.033	0.021	0.966	0.856	432.1	384.9	0.991	0.766	0.087	0.238	0.904	0.528	405.4	236.1
		200	0.998	0.945	0.046	0.022	0.952	0.923	426.4	412.9	0.994	0.882	0.113	0.210	0.881	0.672	396.2	301.7	
		50	0.990	0.745	0.026	0.011	0.964	0.734	431.2	338.4	0.830	0.610	0.026	0.168	0.804	0.442	363.2	202.8	
		75	0.998	0.837	0.030	0.017	0.968	0.820	433.0	370.6	0.986	0.715	0.028	0.206	0.958	0.509	428.4	228.3	
	100	0.998	0.881	0.028	0.017	0.970	0.864	434.0	388.3	0.997	0.780	0.033	0.219	0.964	0.561	431.4	250.9		
	200	0.999	0.945	0.033	0.022	0.965	0.923	431.9	413.1	0.998	0.893	0.035	0.201	0.962	0.692	430.7	310.7		
	50	0.985	0.713	0.041	0.018	0.944	0.695	509.3	414.5	0.805	0.510	0.098	0.173	0.707	0.337	387.3	192.6		
	75	0.994	0.804	0.056	0.024	0.938	0.780	501.9	446.9	0.973	0.592	0.135	0.207	0.839	0.386	439.6	211.1		
	100	0.995	0.851	0.066	0.025	0.929	0.826	495.5	465.5	0.992	0.654	0.147	0.217	0.845	0.437	440.6	234.7		
	200	0.996	0.923	0.093	0.028	0.903	0.894	477.1	492.4	0.997	0.783	0.161	0.212	0.836	0.570	434.7	300.2		
	50	0.986	0.727	0.037	0.014	0.948	0.713	512.2	424.3	0.812	0.526	0.069	0.160	0.742	0.365	412.0	209.5		
	75	0.995	0.818	0.046	0.017	0.948	0.801	509.5	458.9	0.973	0.611	0.096	0.188	0.877	0.422	464.8	232.1		
	100	0.995	0.864	0.055	0.019	0.940	0.845	503.5	475.8	0.991	0.680	0.107	0.201	0.884	0.478	465.6	257.2		
	200	0.996	0.931	0.069	0.022	0.927	0.909	493.6	501.2	0.996	0.813	0.120	0.200	0.877	0.613	459.9	322.4		
	50	0.984	0.707	0.033	0.015	0.951	0.692	586.3	485.2	0.801	0.467	0.075	0.144	0.726	0.323	455.4	211.8		
	75	0.993	0.799	0.043	0.017	0.950	0.782	578.9	524.0	0.968	0.557	0.106	0.169	0.862	0.388	505.7	240.5		
	100	0.994	0.846	0.053	0.019	0.941	0.827	568.5	541.4	0.987	0.620	0.112	0.176	0.876	0.444	509.7	268.5		
	200	0.995	0.918	0.077	0.024	0.919	0.894	545.8	566.8	0.994	0.751	0.128	0.184	0.866	0.567	499.2	331.2		
	50	0.984	0.723	0.032	0.011	0.953	0.712	587.7	498.3	0.808	0.493	0.050	0.131	0.758	0.362	486.0	239.2		
	75	0.993	0.815	0.039	0.013	0.955	0.803	584.0	536.6	0.972	0.581	0.070	0.157	0.902	0.425	541.2	264.2		
	100	0.994	0.861	0.043	0.016	0.951	0.845	579.5	552.2	0.989	0.643	0.079	0.173	0.910	0.470	540.8	283.5		
	200	0.995	0.928	0.056	0.020	0.939	0.907	566.4	575.9	0.996	0.791	0.089	0.178	0.907	0.612	534.4	356.2		

Table 3: Multiple Clubs: Success Percentages, $N = \{10, 20, 30, 40\}$, 5% significance level

Data Type					No Constant		With Constant		Data Type					No Constant		With Constant	
N	k	m	ρ	T	MCL	HF	MCL	HF	N	k	m	ρ	T	MCL	HF	MCL	HF
10	2	4,4	0.2	50	90.1%	54.3%	28.8%	28.0%	30	5	6,6,6,5,5	0.2	50	79.3%	8.2%	1.2%	3.4%
				75	93.0%	68.4%	80.8%	40.0%					75	89.4%	23.7%	50.6%	11.3%
				100	92.9%	74.0%	83.6%	47.0%					100	88.1%	36.1%	55.3%	18.5%
				200	92.2%	83.6%	80.9%	62.6%					200	77.8%	57.9%	44.1%	37.5%
			0.6	50	87.9%	56.5%	32.0%	34.1%				0.6	50	78.3%	8.0%	1.7%	5.1%
				75	92.0%	69.6%	89.8%	44.1%					75	90.1%	24.2%	71.0%	13.6%
	3	3,3,2	0.2	50	86.1%	63.0%	25.2%	27.6%		6	5,5,5,5,4,4	0.2	50	81.1%	12.3%	2.0%	4.3%
				75	86.9%	71.6%	45.5%	38.7%					75	88.6%	29.2%	49.8%	12.9%
				100	85.4%	76.1%	41.2%	45.7%					100	87.3%	39.1%	51.3%	20.4%
				200	81.0%	83.7%	33.2%	63.1%					200	82.6%	59.9%	39.5%	37.7%
			0.6	50	86.5%	63.8%	36.4%	35.1%				0.6	50	80.8%	12.4%	2.9%	6.7%
				75	89.3%	72.8%	71.2%	45.0%					75	89.6%	29.7%	72.4%	16.8%
20	4	5,5,4,4	0.2	50	82.7%	27.5%	5.5%	31.9%	40	7	6,6,6,6,6,4,4	0.2	50	74.5%	1.5%	0.4%	6.9%
				75	88.5%	46.1%	49.1%	44.3%					75	88.3%	9.5%	46.0%	22.0%
				100	87.1%	56.1%	49.3%	51.3%					100	88.0%	20.2%	54.8%	33.4%
				200	83.2%	72.2%	39.8%	65.9%					200	84.2%	43.7%	44.0%	55.2%
			0.6	50	82.2%	27.9%	7.2%	35.0%				0.6	50	72.5%	1.3%	0.3%	7.4%
				75	90.2%	45.7%	73.9%	47.1%					75	89.8%	9.7%	63.8%	23.5%
	5	4,4,4,3,3	0.2	50	84.3%	26.1%	8.9%	31.3%		8	5,5,5,5,5,5,4,4	0.2	50	77.8%	3.6%	0.9%	10.2%
				75	88.1%	43.2%	50.6%	42.8%					75	88.8%	15.6%	47.5%	27.0%
				100	87.0%	52.9%	48.7%	51.1%					100	87.0%	26.8%	51.4%	39.1%
				200	82.6%	68.4%	38.9%	64.8%					200	83.0%	49.9%	39.7%	58.5%
			0.6	50	83.5%	26.1%	12.0%	34.1%				0.6	50	77.2%	3.2%	1.0%	11.4%
				75	89.3%	44.4%	75.0%	45.0%					75	89.2%	15.4%	68.8%	29.4%
				100	89.9%	53.6%	80.0%	53.1%					100	90.0%	27.1%	80.7%	40.4%
				200	88.6%	68.8%	76.5%	66.3%					200	88.7%	50.4%	69.4%	58.8%

Table 4: Multiple Clubs: Success Counts for Poisson settlement, $N = \{10, 20, 30, 40\}$, 5% significance level

Data Type					No Constant		With Constant		Data Type					No Constant		With Constant	
N	k	m	ρ	T	MCL	HF	MCL	HF	N	k	m	ρ	T	MCL	HF	MCL	HF
10	2	6,3	0.2	50	92.4%	49.7%	34.4%	27.5%	30	5	6,6,4,4,3	0.2	50	79.6%	15.2%	2.4%	4.4%
				75	96.3%	66.8%	86.0%	37.5%					75	85.7%	33.4%	27.2%	11.7%
				100	93.5%	73.6%	86.2%	45.2%					100	83.1%	46.4%	24.7%	19.7%
				200	93.6%	83.1%	85.8%	60.7%					200	78.1%	64.5%	20.0%	36.9%
			0.6	50	91.3%	50.2%	37.8%	31.6%				50	80.0%	16.5%	4.7%	6.9%	
				75	94.2%	67.7%	92.2%	41.0%				75	87.3%	34.9%	57.5%	16.1%	
	3	4,4,2	0.2	75	92.5%	74.5%	93.7%	47.9%		100	87.5%	45.9%	59.3%	25.3%			
				100	92.5%	74.5%	93.7%	47.9%		100	87.5%	45.9%	59.3%	25.3%			
				200	94.2%	83.3%	95.5%	61.7%		200	85.8%	65.6%	56.8%	41.0%			
				0.6	50	91.3%	53.4%	31.8%		33.1%	50	78.7%	5.8%	1.2%	3.6%		
			75		95.5%	69.0%	78.7%	40.1%		75	93.4%	20.4%	65.3%	11.5%			
			0.6	100	92.0%	72.5%	80.6%	47.4%		100	92.2%	32.9%	76.7%	20.8%			
200	93.5%	81.4%		81.4%	63.6%	200	92.0%	56.5%	80.7%	40.7%							
20	4	5,5,3,2	0.2	50	90.4%	53.9%	37.7%	36.5%	40	7	7,7,7,6,6,4,2	0.2	50	74.6%	1.9%	0.2%	1.0%
				75	93.9%	69.6%	89.8%	44.5%					75	88.9%	11.0%	47.3%	5.1%
				100	91.8%	73.2%	92.7%	49.7%					100	88.0%	22.7%	59.4%	10.7%
				200	93.6%	81.1%	94.8%	64.1%					200	85.5%	46.4%	51.8%	27.8%
			0.6	50	83.2%	35.2%	13.7%	15.4%				50	73.4%	2.0%	0.3%	1.4%	
				75	89.5%	53.3%	72.4%	26.3%				75	90.3%	11.0%	65.6%	6.8%	
	5	4,4,4,3,2	0.2	100	88.5%	61.0%	76.4%	33.7%		100	89.5%	22.1%	83.2%	13.1%			
				200	89.8%	74.6%	76.7%	49.9%		200	90.5%	46.0%	84.7%	30.8%			
				0.6	50	87.4%	35.4%	13.4%		12.9%	50	24.7%	7.8%	0.1%	0.2%		
					75	89.3%	50.6%	54.9%		21.3%	75	17.2%	20.1%	0.1%	0.7%		
			0.6	100	87.7%	59.5%	52.7%	28.8%		100	10.9%	28.6%	0.0%	1.6%			
				200	85.1%	72.5%	48.3%	44.1%		200	1.3%	47.4%	0.0%	8.5%			
8	6,6,4,4,3,3,2	0.2	50	84.8%	34.6%	18.7%	18.0%	50	41.0%	8.8%	0.3%	0.6%					
			75	90.1%	50.5%	80.3%	27.0%	75	35.9%	22.6%	0.4%	2.1%					
			100	89.6%	59.8%	83.1%	34.9%	100	30.1%	32.4%	0.1%	4.2%					
			200	90.9%	72.4%	84.1%	49.4%	200	12.2%	51.4%	0.0%	17.2%					

Table 5: Country Groups based on Economic Characteristics and Data Availability for Growth Application

Penn (T=65, N=55)	Argentina, Australia, Austria, Belgium, Bolivia, Brazil, Canada, Colombia, Costa.Rica, Cyprus, Denmark, Congo , Ecuador, Egypt, El.Salvador, Ethiopia, Finland, France, Germany, Guatemala, Honduras, Iceland, India, Ireland, Israel, Italy, Japan, Kenya, Luxembourg, Mauritius, Mexico, Morocco, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Pakistan, Panama, Peru, Philippines, Portugal, South.Africa, Spain, Sri Lanka, Sweden, Switzerland, Thailand, Trinidad and Tobago, Turkey, Uganda, United Kingdom, United States, Uruguay, Venezuela
Maddison (T=61, N=95)	Albania, Algeria, Angola, Argentina, Australia, Austria, Bahrain, Bangladesh, Belgium, Bolivia, Brazil, Bulgaria, Burkina Faso, Burma, Cambodia, Cameroon, Canada, Chile, China, Colombia, Congo Kinshasa, Costa Rica, Cote d'Ivoire, Czecho-slovakia, Denmark, Dominican Rep., Ecuador, Egypt, Ethiopia, Finland, France, Germany, Ghana, Greece, Guatemala, Hong Kong, Hungary, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Kuwait, Madagascar, Malawi, Malaysia, Mali, Mexico, Morocco, Mozambique, Netherlands, New Zealand, Niger, Nigeria, Norway, Oman, Pakistan, Peru, Philippines, Poland, Portugal, Qatar, Romania, Saudi Arabia, Senegal, Singapore, South Africa, South Korea, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Taiwan, Tanzania, Thailand, Tunisia, Turkey, UAE, Uganda, United Kingdom, United States, Uruguay, Venezuela, Vietnam, Yemen, Zambia, Zimbabwe
1930 (T=81, N=36)	Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, Colombia, Costa Rica, Denmark, Ecuador, Finland, France, Germany, Greece, Guatemala, India, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Peru, Portugal, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Turkey, UK, Uruguay, USA, Venezuela
Europe (T=61, N=22)	Albania, Austria, Belgium, Bulgaria, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Hungary, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Romania, Sweden, Turkey
G7 (T=61, N=7)	Canada, France, Germany, Italy, Japan, UK, USA
S&P (T=61, N=19)	Brazil, Chile, China, Colombia, Egypt, Greece, Hungary, India, Indonesia, Malaysia, Mexico, Morocco, Peru, Philippines, Poland, South Africa, Taiwan, Thailand, Turkey

Table 6: Application on Growth Convergence: 5% significance level

Data / T / N	Model	# 2	# 3	# 4	# 5	# 6	# 7
Penn T=65,N=55	MCL	26	90	7			
	HF	6	8	1	3		
Maddison T=61,N=95	MCL	60	69	1	1		
	HF	18	7	6	0	0	2
1930 T=81,N=36	MCL	14	21	7			
	HF	5	2	5			
Europe T=61,N=22	MCL	17					
	HF	4	3	0	1		
Europe+G7 T=61,N=25	MCL	25	5				
	HF	5	2	1	1		
Europe+S&P T=61,N=37	MCL	11	14				
	HF	7	3	1	2		
G7+S&P T=61,N=26	MCL	10	9				
	HF	5	4	1			

Figure 1: A sample undirected graph

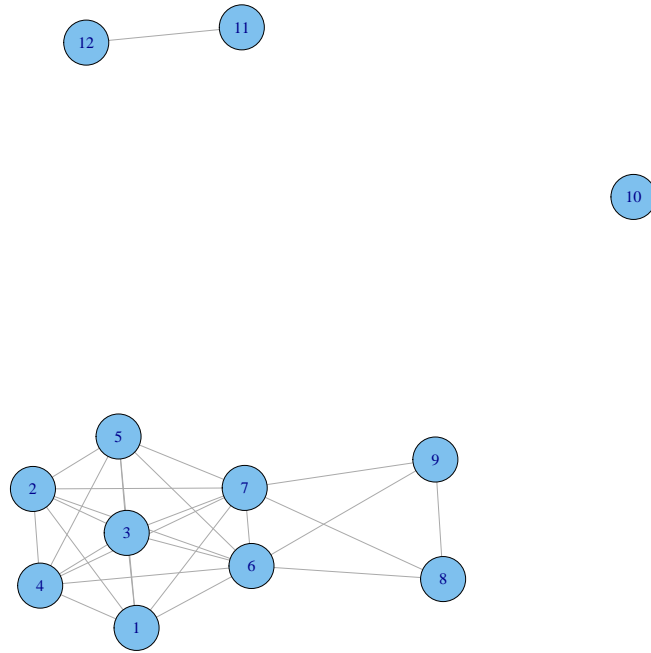


Figure 2: A sample maximum clique

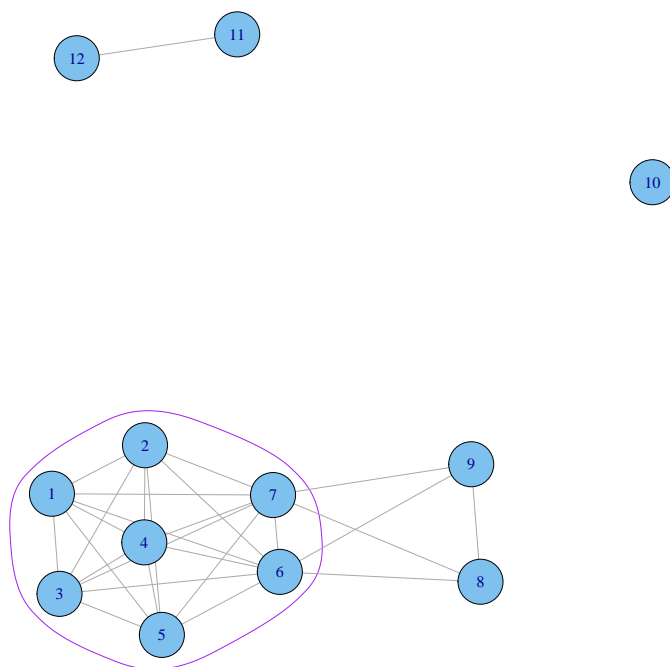


Figure 3: An example of a broken club

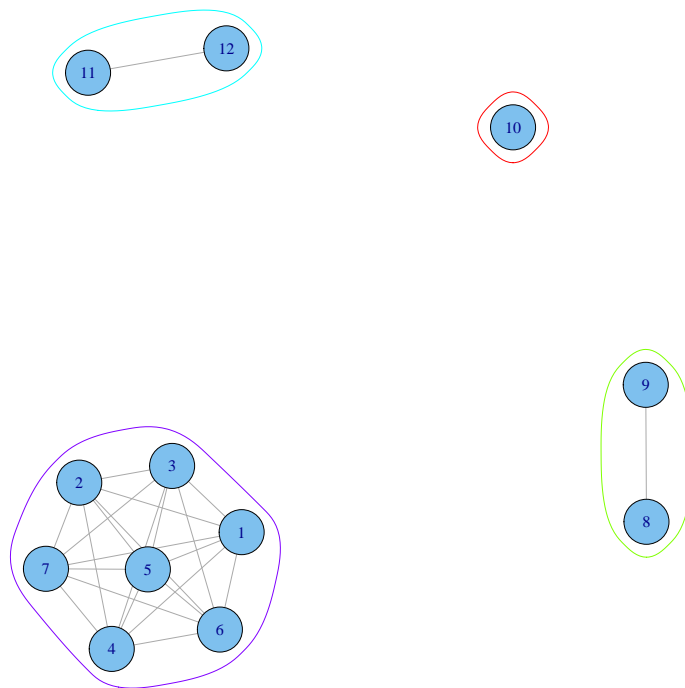


Figure 4: A sample maximal clique

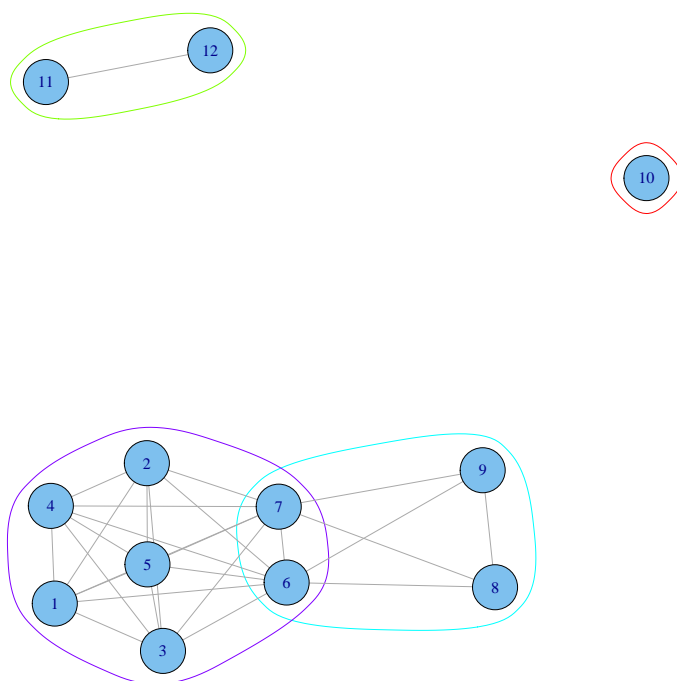




Figure 5: Illustration of a Club: Europe + G7 (MCL)



Figure 6: Illustration of a Club: Europe + G7 (MCL)

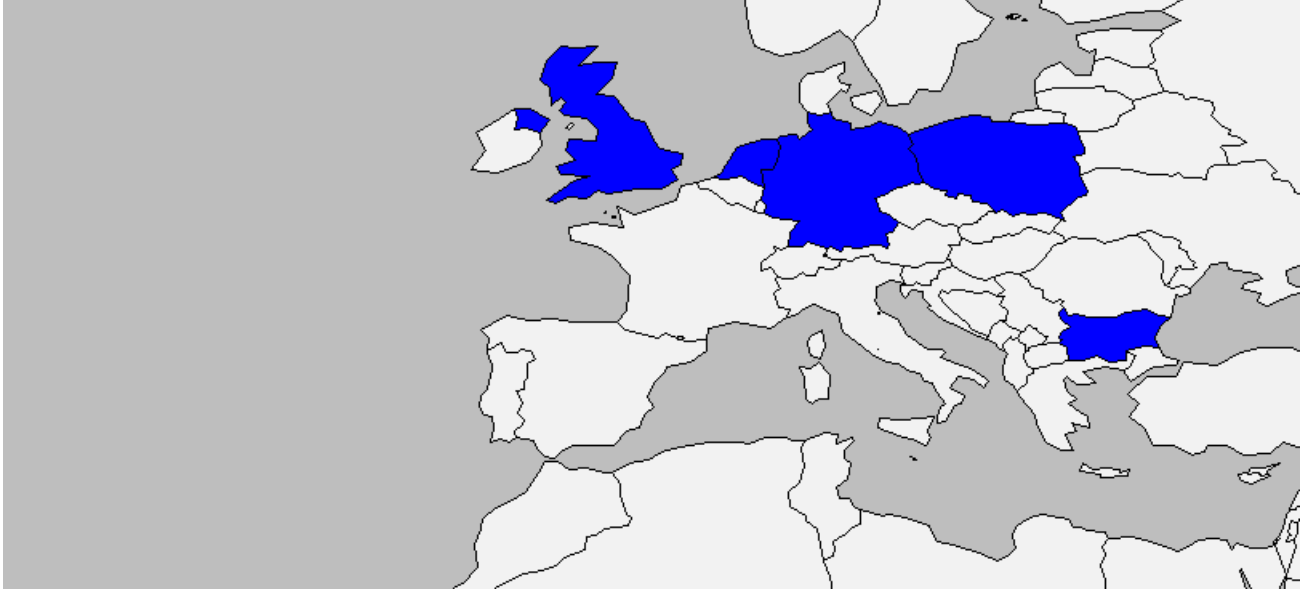


Figure 7: Illustration of a Club: Europe + G7 (HF)



Figure 8: Illustration of a Club: Europe + S&P (MCL)



Figure 9: Illustration of a Club: Europe + S&P (HF)



Figure 10: Illustration of a Club: G7 + S&P (MCL)



Figure 11: Illustration of a Club: G7 + S&P (MCL)



Figure 12: Illustration of a Club: G7 + S&P (HF)

References

- Abbott, A. and G. De Vita (2013). Testing for long-run convergence across regional house prices in the UK: a pairwise approach. *Applied Economics* 45(10), 1227–1238.
- Abbott, A., G. De Vita, and L. Altinay (2012). Revisiting the convergence hypothesis for tourism markets: Evidence from turkey using the pairwise approach. *Tourism Management* 33(3), 537–544.
- Apergis, N. and P. Padhi (2013). Health expenses and economic growth: convergence dynamics across the Indian States. *International journal of health care finance and economics* 13(3-4), 261–277.
- Azariadis, C. (1996). The economics of poverty traps part one: complete markets. *Journal of economic growth* 1(4), 449–486.
- Azariadis, C. and A. Drazen (1990). Threshold externalities in economic development. *The Quarterly Journal of Economics* 105(2), 501–526.
- Baumol, W. J. (1986). Productivity growth, convergence, and welfare: what the long-run data show. *The American Economic Review*, 1072–1085.
- Bernard, A. and S. Durlauf (1996). Interpreting tests of the convergence hypothesis. *Journal of Econometrics* 71, 161–173.
- Bernard, A. B. and S. N. Durlauf (1995). Convergence in international output. *Journal of applied econometrics* 10(2), 97–108.
- Beylunioglu, F. C., T. Stengos, and E. Yazgan (2018). Regime switching with structural breaks in output convergence. *Studies in Nonlinear Dynamics and Econometrics (forthcoming)*.
- Bron, C. and J. Kerbosch (1973). Algorithm 457: finding all cliques of an undirected graph. *Communications of the ACM* 16(9), 575–577.
- Chatterji, M. (1992). Convergence clubs and endogenous growth. *Oxford Review of economic policy* 8(4), 57–69.

- Corrado, L., R. Martin, and M. Weeks (2005). Identifying and interpreting regional convergence clusters across europe*. *The Economic Journal* 115(502), C133–C160.
- Dufrénot, G., V. Mignon, and T. Naccache (2012). Testing catching-up between the developing countries: “growth resistance” and sometimes “growth tragedy”. *Bulletin of Economic Research* 64(4), 470–508.
- Durlauf, S. N. and P. A. Johnson (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10(4), 365–384.
- Durlauf, S. N., P. A. Johnson, and J. R. Temple (2005). Growth econometrics. *Handbook of economic growth* 1, 555–677.
- Fritsche, U. and V. Kuzin (2011). Analysing convergence in europe using the non-linear single factor model. *Empirical Economics* 41(2), 343–369.
- Galor, O. (1996). Convergence? inferences from theoretical models. *The Economic Journal*, 1056–1069.
- Galor, O. (2011). *Unified growth theory*. Princeton University Press.
- Galor, O. and O. Moav (2002). Natural selection and the origin of economic growth. *The Quarterly Journal of Economics* 117(4), 1133–1191.
- Galor, O. and D. N. Weil (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American economic review* 90(4), 806–828.
- Hausmann, R., L. Pritchett, and D. Rodrik (2005). Growth accelerations. *Journal of Economic Growth* 10(4), 303–329.
- Hobijn, B. and P. H. Franses (2000). Asymptotically perfect and relative convergence of productivity. *Journal of Applied Econometrics* 15(1), 59–81.
- Ikeno, H. (2014). Pairwise tests of convergence of japanese local price levels. *International Review of Economics & Finance* 31, 232–248.
- Johnson, P. and C. Papageorgiou (2017). What remains of cross-country convergence? *Journal of Economic Literature* (forthcoming).

- Kim, Y. S. and J. J. Rous (2012). House price convergence: Evidence from us state and metropolitan area panels. *Journal of Housing Economics* 21(2), 169–186.
- Konc, J. and D. Janezic (2007). An improved branch and bound algorithm for the maximum clique problem. *proteins* 4, 5.
- Özkan, H., T. Stengos, and M. E. Yazgan (2014). Persistence in convergence and club formation.
- Pesaran, H. M. (2007). A pair-wise approach to testing for output and growth convergence. *Journal of Econometrics* 138(1), 312–355.
- Pesaran, M. and A. Timmermann (1992). A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics* 10, 461–465.
- Phillips, P. C. and D. Sul (2007). Transition modeling and econometric convergence tests. *Econometrica* 75(6), 1771–1855.
- Quah, D. T. (1996). Empirics for economic growth and convergence. *European economic review* 40(6), 1353–1375.
- Stengos, T. and M. E. Yazgan (2014). Persistence in convergence. *Macroeconomic Dynamics* 18(04), 753–782.
- Yilmazkuday, H. (2013). Inflation targeting, flexible exchange rates and inflation convergence. *Applied Economics* 45(5), 593–603.

7 Appendix

Table A3: Multiple Clubs: Success Percentages, $N = \{10, 20, 30, 40\}$, 1% and 10% significance levels

Data Type			No Constant			With Constant							
N	k	m	T	MCL	HF	MCL	HF	HF					
2	4,4	0.2	50	52.0%	88.2%	54.3%	54.1%	1.5%	51.4%	28.0%	28.0%		
			75	96.7%	87.5%	68.4%	68.1%	52.8%	74.1%	40.0%	39.9%	40.0%	
		100	98.3%	87.5%	74.0%	73.6%	91.3%	72.8%	47.0%	46.8%	47.0%	46.8%	
		200	97.8%	86.4%	83.6%	82.9%	93.2%	70.4%	62.6%	62.3%	62.6%	62.3%	
		50	52.7%	85.6%	56.5%	56.3%	1.6%	59.4%	34.1%	34.1%	34.1%	34.1%	
		75	95.8%	86.8%	69.6%	69.3%	54.8%	87.9%	44.1%	44.0%	44.1%	44.0%	
	100	97.7%	86.9%	75.0%	74.6%	94.7%	88.6%	50.7%	50.5%	50.7%	50.5%		
	200	97.8%	88.0%	83.6%	83.1%	99.0%	89.3%	63.4%	63.0%	63.4%	63.0%		
	10		0.2	50	60.9%	81.6%	63.0%	62.1%	4.0%	29.6%	27.6%	27.5%	27.6%
				75	93.0%	80.4%	71.6%	69.8%	45.6%	33.3%	38.7%	38.5%	38.7%
			100	93.5%	78.7%	76.1%	73.7%	63.4%	29.0%	45.7%	45.2%	45.7%	45.2%
			200	90.2%	73.5%	83.7%	80.7%	53.0%	23.3%	63.1%	62.1%	63.1%	62.1%
50			61.7%	82.5%	63.8%	62.6%	4.9%	50.1%	35.1%	35.0%	35.1%	35.0%	
75			94.7%	83.1%	72.8%	70.9%	57.9%	62.7%	45.0%	44.8%	45.0%	44.8%	
100	96.4%	83.5%	76.9%	74.8%	86.9%	62.4%	51.6%	51.2%	51.6%	51.2%			
200	95.4%	81.6%	83.8%	81.0%	86.2%	58.5%	64.9%	63.7%	64.9%	63.7%			
3	3,3,2	0.2	50	23.5%	82.9%	27.5%	27.2%	0.0%	20.2%	31.9%	31.8%	31.9%	
			75	92.8%	82.0%	46.1%	45.2%	21.8%	40.2%	44.3%	44.3%	44.3%	44.3%
		100	94.9%	80.3%	56.1%	54.9%	68.3%	36.7%	51.3%	51.3%	51.3%	51.3%	
		200	91.9%	76.0%	72.2%	70.2%	60.0%	28.9%	65.9%	65.8%	65.9%	65.8%	
		50	23.3%	82.4%	27.9%	27.5%	0.0%	31.4%	35.0%	35.0%	35.0%	35.0%	
		75	93.5%	83.8%	45.7%	44.8%	25.4%	69.6%	47.1%	47.1%	47.1%	47.1%	
	100	96.7%	83.5%	56.5%	55.2%	86.0%	69.6%	53.9%	53.8%	53.9%	53.8%		
	200	96.4%	83.3%	71.5%	69.4%	91.4%	66.0%	67.1%	67.0%	67.1%	67.0%		
	4	5,5,4,4	0.2	50	31.4%	82.4%	26.1%	25.5%	0.0%	23.6%	31.3%	31.3%	31.3%
				75	92.7%	81.6%	43.2%	41.9%	27.3%	40.7%	42.8%	42.7%	42.8%
			100	94.7%	80.4%	52.9%	51.0%	68.6%	35.9%	51.1%	51.0%	51.1%	51.0%
			200	91.5%	75.6%	68.4%	64.9%	59.9%	27.8%	64.8%	64.6%	64.8%	64.6%
50			31.7%	81.9%	26.1%	25.5%	0.1%	35.6%	34.1%	34.1%	34.1%	34.1%	
75			93.5%	83.0%	44.4%	43.0%	33.5%	69.4%	45.0%	45.0%	45.0%	45.0%	
100	96.6%	83.8%	53.6%	51.8%	86.2%	69.1%	53.1%	53.0%	53.1%	53.0%			
200	95.9%	81.7%	68.8%	65.9%	90.4%	65.0%	66.3%	66.2%	66.3%	66.2%			
40	5,5,5,5,5,4,4	0.2	50	8.0%	81.5%	3.6%	3.5%	0.0%	10.4%	10.2%	10.2%	10.2%	
			75	90.4%	82.4%	15.6%	15.1%	7.5%	40.9%	27.0%	27.0%	27.0%	
		100	94.9%	80.2%	26.8%	25.4%	65.7%	37.1%	39.1%	39.0%	39.1%	39.0%	
		200	92.0%	75.8%	49.9%	46.9%	60.5%	28.3%	58.5%	58.4%	58.5%	58.4%	
		50	7.6%	81.0%	3.2%	3.0%	0.0%	14.9%	11.4%	11.4%	11.4%	11.4%	
		75	91.0%	83.3%	15.4%	14.8%	9.9%	69.2%	29.4%	29.3%	29.4%	29.3%	
	100	96.7%	83.5%	27.1%	26.0%	80.9%	69.8%	40.4%	40.4%	40.4%	40.4%		
	200	96.0%	82.3%	50.4%	47.3%	92.1%	64.3%	58.8%	58.7%	58.8%	58.7%		
	5	6,6,6,6,6,4,4	0.2	50	2.9%	80.3%	1.3%	1.3%	0.0%	9.5%	7.4%	7.4%	7.4%
				75	87.6%	83.9%	9.7%	9.6%	4.5%	69.6%	23.5%	23.4%	23.5%
			100	96.6%	83.9%	20.3%	19.8%	74.5%	73.0%	34.5%	34.5%	34.5%	34.5%
			200	96.7%	83.9%	45.1%	43.6%	92.4%	69.6%	55.4%	55.3%	55.4%	55.3%
50			8.0%	81.5%	3.6%	3.5%	0.0%	10.4%	10.2%	10.2%	10.2%	10.2%	
75			90.4%	82.4%	15.6%	15.1%	7.5%	40.9%	27.0%	27.0%	27.0%	27.0%	
100	94.9%	80.2%	26.8%	25.4%	65.7%	37.1%	39.1%	39.0%	39.1%	39.0%			
200	92.0%	75.8%	49.9%	46.9%	60.5%	28.3%	58.5%	58.4%	58.5%	58.4%			

Table A4: Multiple Clubs: Success Counts for Poisson settlement, $N = \{10, 20, 30, 40\}$, 1000 replications, 1% and 10% significance levels

Data Type			No Constant			With Constant			Data Type			No Constant			With Constant												
N	k	m	T	ρ	MCL	HF	MCL	HF	T	ρ	MCL	HF	MCL	HF	T	ρ	MCL	HF									
2		6,3	50	0.2	59.5%	90.3%	49.7%	49.4%	1.7%	56.3%	27.5%	27.5%	56.3%	27.5%	27.5%	50	21.8%	78.9%	15.2%	15.1%	0.0%	10.5%	1%	10%	4.4%	4.3%	
					75	98.4%	92.7%	66.8%	66.4%	64.2%	78.0%	37.5%	37.3%	64.2%	78.0%	33.4%	33.0%	14.9%	19.0%	11.7%	11.7%						
					100	98.7%	87.9%	73.6%	72.7%	93.8%	78.2%	45.2%	45.1%	93.8%	78.2%	46.4%	45.6%	45.7%	16.4%	19.7%	19.5%						
					200	98.6%	88.3%	83.1%	82.1%	95.0%	76.1%	60.7%	60.1%	95.0%	76.1%	64.5%	63.0%	36.0%	12.4%	36.9%	36.4%						
					50	59.0%	88.8%	50.2%	49.9%	1.7%	66.6%	31.6%	31.6%	1.7%	66.6%	31.6%	31.6%	0.0%	22.2%	6.9%	6.8%						
					75	97.2%	90.9%	67.7%	67.3%	65.5%	88.8%	41.0%	40.8%	65.5%	88.8%	41.0%	40.8%	21.8%	47.9%	16.1%	16.0%						
10			50	0.6	97.8%	88.7%	74.5%	73.6%	96.9%	88.6%	47.9%	47.7%	96.9%	88.6%	45.9%	45.3%	25.3%	25.1%	100	94.6%	81.4%	34.9%	34.4%	72.7%	45.3%	25.3%	25.1%
					200	98.7%	89.3%	83.3%	82.3%	99.4%	91.5%	61.7%	61.1%	99.4%	91.5%	65.6%	63.9%	76.7%	45.2%	41.0%	40.5%						
					50	58.3%	88.0%	53.4%	53.0%	2.0%	50.9%	33.1%	33.1%	2.0%	50.9%	33.1%	33.1%	0.0%	15.1%	3.6%	3.6%						
					75	98.2%	90.8%	69.0%	67.9%	62.0%	70.5%	40.1%	39.9%	62.0%	70.5%	40.1%	39.9%	13.0%	62.2%	11.5%	11.4%						
					100	98.7%	87.0%	72.5%	71.8%	91.1%	71.5%	47.4%	47.2%	91.1%	71.5%	32.9%	32.1%	76.8%	64.4%	20.8%	20.4%						
					200	97.9%	86.3%	81.4%	79.7%	93.2%	71.5%	63.6%	63.2%	93.2%	71.5%	56.5%	54.5%	92.6%	71.1%	40.7%	39.7%						
3		4,4,2	50	0.6	57.3%	86.7%	53.9%	53.4%	2.1%	63.7%	36.5%	36.5%	2.1%	63.7%	36.5%	36.5%	0.0%	19.8%	4.8%	4.8%							
					75	97.3%	89.7%	69.6%	68.5%	64.8%	86.8%	44.5%	44.1%	64.8%	86.8%	44.5%	44.1%	13.1%	79.9%	12.1%	12.0%						
					100	97.5%	87.8%	73.2%	72.5%	96.3%	86.7%	49.7%	49.5%	96.3%	86.7%	49.7%	49.5%	82.0%	83.6%	23.5%	23.0%						
					200	98.0%	88.3%	81.1%	79.4%	98.9%	90.2%	64.1%	63.6%	98.9%	90.2%	56.2%	54.2%	98.9%	88.9%	41.2%	40.3%						
					50	37.3%	81.4%	34.6%	34.1%	0.1%	17.4%	9.2%	9.2%	0.1%	17.4%	9.2%	9.2%	0.0%	5.3%	1.0%	1.0%						
					75	92.5%	81.2%	51.3%	50.0%	32.8%	33.0%	19.2%	19.0%	32.8%	33.0%	11.0%	10.9%	2.6%	46.4%	5.1%	5.1%						
4		5,5,3,2	50	0.2	94.4%	79.6%	61.2%	59.6%	61.8%	30.4%	27.2%	27.0%	61.8%	30.4%	22.7%	22.2%	65.1%	47.3%	10.7%	10.5%							
					200	92.4%	75.9%	73.0%	70.2%	58.1%	28.5%	46.2%	45.5%	58.1%	28.5%	46.4%	44.9%	71.7%	38.6%	27.8%	27.4%						
					50	37.2%	79.9%	35.2%	34.5%	0.0%	36.9%	15.4%	15.4%	0.0%	36.9%	15.4%	15.4%	0.0%	8.7%	1.4%	1.4%						
					75	93.7%	84.6%	53.3%	52.2%	42.3%	63.8%	26.3%	26.1%	42.3%	63.8%	26.3%	26.1%	3.3%	70.4%	6.8%	6.7%						
					100	96.0%	82.6%	61.0%	59.4%	86.4%	64.2%	33.7%	33.3%	86.4%	64.2%	22.1%	21.7%	76.1%	74.8%	13.1%	13.0%						
					200	96.7%	82.8%	74.6%	71.9%	90.6%	65.5%	49.9%	49.2%	90.6%	65.5%	46.0%	44.5%	95.4%	75.2%	30.8%	30.2%						
20			50	0.6	40.0%	83.3%	35.4%	34.9%	0.3%	25.3%	12.9%	12.9%	0.3%	25.3%	12.9%	12.9%	0.0%	0.2%	0.2%								
					75	95.1%	83.4%	50.6%	49.2%	40.2%	44.0%	21.3%	21.1%	40.2%	44.0%	20.1%	19.2%	0.4%	0.0%	0.7%	0.6%						
					100	95.4%	81.2%	59.5%	57.9%	73.2%	40.8%	28.8%	28.4%	73.2%	40.8%	28.6%	27.2%	0.3%	0.0%	1.6%	1.6%						
					200	94.5%	77.4%	72.5%	69.7%	68.8%	37.4%	44.1%	43.1%	72.5%	69.7%	47.4%	44.6%	0.0%	0.0%	8.5%	8.2%						
					50	39.8%	81.7%	34.6%	34.1%	0.3%	43.4%	18.0%	18.0%	0.3%	43.4%	18.0%	18.0%	0.0%	0.7%	0.6%	0.6%						
					75	95.3%	85.6%	50.5%	49.3%	46.3%	73.8%	27.0%	26.9%	46.3%	73.8%	22.6%	21.6%	0.9%	0.1%	2.1%	2.1%						
40			50	0.6	96.7%	83.1%	59.8%	57.9%	89.9%	71.9%	34.9%	34.4%	89.9%	71.9%	32.4%	30.8%	0.8%	0.0%	4.2%	4.1%							
					200	97.0%	84.7%	72.4%	69.8%	94.9%	74.7%	49.4%	48.2%	94.9%	74.7%	51.4%	47.9%	0.0%	0.0%	17.2%	16.6%						