The technology clubs: The distribution of knowledge across nations

Fulvio Castellacci [a,*], Daniele Archibugi [b,c]

[a] Department of International Economics, Norwegian Institute of International Affairs (NUPI), POB 8159, Dep. 0033 Oslo, Norway
[b] Italian National Research Council, via dei Taurini 19, 00185 Rome, Italy
[c] Department of Management, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

ARTICLE INFO

ABSTRACT

The convergence clubs literature in applied growth theory suggests that countries that differ in terms of structural characteristics and initial conditions tend to experience diverging growth performances. What is the role of technological knowledge for the formation of clubs? The paper investigates this unexplored question by carrying out an empirical study of the cross-country distribution of knowledge in a large sample of developed and developing economies in the 1990s. The results indicate the existence of three technology clubs characterized by markedly different levels of development. The clubs also differ with respect to the dynamics of their capabilities over the decade, as the most advanced group and the intermediate one are found to be much more dynamic than the large cluster of less developed economies.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays it is well accepted that the distribution of knowledge across nations is unbalanced. Countries in the world economy are characterized by different levels of technological development and have unequal access to knowledge stocks, and this is a major factor to explain their different competitive patterns and diverging economic dynamics in the long run (for an historical overview, see von Tunzelmann, 1995).

The investigation of the reasons behind cross-country differences in the creation and exploitation of knowledge constitutes, in a nutshell, the main interest of the national innovation systems literature (Lundvall, 1992; Nelson, 1993; Edquist, 1997). Studies in this tradition have become increasingly popular in the last couple of decades, and have greatly enriched our understanding of the functioning and evolution of the competencies of nations (Fagerberg and Verspagen, 2002). Besides the great variety of theoretical

insights that the systemic perspective has made it possible to achieve, several empirically oriented studies have also tried to operationalize the approach by measuring different aspects of the process of knowledge creation and dissemination in a large sample of countries, and by investigating the extent of cross-country differences (Archibugi and Coco, 2004a,b; Grupp and Mogee, 2004).

To what extent have these new insights been recognised, not to say accepted, by applied growth theory in mainstream economics? Not much, arguably. Applied growth theory has achieved considerable progress in the last decade, and, moving away from the traditional cross-country regression framework, it has made use of a wide array of econometric techniques to point to the existence of a great deal of heterogeneity across countries.[1] However, the study of macroeconomic variety has mainly been analysed in terms of economic factors, while the role of cross-country differences in the ability to create and imitate advanced knowledge has not been investigated to the same extent.

---

* Corresponding author. Tel.: +47 22994000; fax: +47 22362182.
E-mail address: fc@nupi.no (F. Castellacci).

---

[1] See overviews by Temple (1999), Islam (2003) and Castellacci (2007).

This is particularly evident when we look at the recent flourishing of empirical studies in the *convergence clubs* tradition (Durlauf and Johnson, 1995). The general idea of this literature, that countries that differ in terms of structural characteristics and initial conditions tend to experience diverging growth performances, has increasingly attracted the attention of applied growth scholars, who have recently proposed a range of different economic explanations for the existence of convergence clubs in the world economy (Desdoigts, 1999; Hobijn and Franses, 2000; Johnson and Takeyama, 2001; Papageorgiou, 2002; Fiaschi and Lavezzi, 2003; Canova, 2004). None of these previous works, however, has investigated the uneven distribution of knowledge across nations and the importance of this for the formation of clubs.

For innovation scholars, the idea that cross-country differences in knowledge stocks and technological capabilities may be one major factor to explain the existence of different country clubs is a natural hypothesis to investigate (see Godinho et al., 2005). What is the role of technological knowledge for the formation of clubs? The paper investigates this unexplored question by carrying out an empirical study of cross-country differences in the ability to exploit the process of knowledge creation and dissemination, focusing on a large sample of developed and developing economies in the 1990s.

The purpose of the paper is thus to study the existence, characteristics and dynamics of different *technology clubs* in the world economy. The study, empirical in nature, will explore the existence of groups of countries characterized by different levels of knowledge stocks and technological infrastructures, point out their characteristics and the distance that separates them, and analyse their different dynamics over time.

The analysis makes use of the ArCo database (Archibugi and Coco, 2004a,b) that contains a set of indicators on various aspects of countries' technological activities for a large number of developed and developing economies. The dataset makes it possible to consider various aspects of countries' technological capabilities, such as their ability to create and to imitate advanced knowledge, their technological infrastructures and their education levels.

The empirical work analyses these data by using a three-step methodology. First, the large number of available indicators is reduced to a smaller number of dimensions by means of a factor analysis. Secondly, these main dimensions are used to identify different country groups through clustering techniques. Finally, a multinomial logit test is estimated in order to assess the statistical precision of the resulting club classification.

The paper is organized as follows. Section 2 presents the data and the indicators used to measure countries' knowledge stocks and technological capabilities, and then reports the results of the factor analysis. Section 3 presents the results of a hierarchical cluster analysis, which explores the existence of various groups of countries differing in terms of their levels of technological development. The results of the cluster analysis show the existence of three clubs characterized by strikingly different levels and diverging dynamics of technological knowledge. Section 4 reports the results of the multinomial logit test of our technology club classification, which provides basic support for its validity and classificatory power. Section 5 concludes the paper by discussing the theoretical interpretation of the empirical results and the related policy implications.

2. Data, indicators and discriminatory factors

The process of knowledge creation and dissemination is a complex and multifaceted aspect of economic progress. Countries' technological capabilities are strictly related to a number of different aspects, such as the basic and advanced human capital base, the infrastructures that support industrial production and innovative activities, and the nation's ability to create, imitate and manage a complex pool of advanced technological knowledge.

The complexity of the concept of technological change presents a formidable challenge to its measurement in empirical analyses. In the economics literature on technology and convergence, the standard approach to measure technological change is to rely on an *indirect* measure, the so-called 'productivity residual' or 'total factor productivity' (TFP; see Prescott, 1998; Islam, 1999). In an attempt to provide a richer characterization and a more precise measurement of countries' ability to create and imitate advanced knowledge, this paper follows a different strategy, and makes use of a set of indicators that measure *directly* various relevant aspects of their technological capabilities.

The empirical analysis is based on the ArCo database, a dataset recently constructed by Archibugi and Coco (2004a,b) that includes a set of indicators of knowledge creation and dissemination for a large number of developed and developing economies. The sample used in this paper consists of 131 countries (listed in Appendix A), whose technological capabilities are measured at two different points in time, at the beginning and at the end of the 1990s. The indicators are briefly described as follows.[2]

2.1. Patents

Number of patents registered at the US Patent and Trademark Office per million people (USPTO, 2002). Patents represent a form of codified knowledge generated by profit-seeking firms and organizations.[3] This indicator is hence a measure of the technological innovations generated for commercial purposes.

---

[2] For a more complete description of the dataset and the methodology used to construct it, see Archibugi and Coco (2004a), while the main differences between the ArCo technology index and other similar analyses to measure countries' technological capabilities are described in Archibugi and Coco (2004b).

[3] To account for yearly fluctuations (which might affect the results in small and medium-sized countries), we have considered a four-year moving average for the 1987–1990 and 1997–2000 periods. Patents are a good proxy of commercially exploitable and proprietary technological inventions, but we are well aware that many inventions are not patented, especially among those invented in developing countries. For surveys on patents as internationally comparable indicators, see Pavitt (1988) and Archibugi (1992).

2.2. Scientific articles

Number of scientific articles per million people (NSF, 2000, 2002).[4] Scientific literature is another important source of codified knowledge. It represents the knowledge generated in the public sector, and most notably in universities and other public research institutes, although researchers working in the business sector publish a significant and growing share of scientific articles.

2.3. Internet penetration

Number of Internet users per 1000 people (ITU, 2001; World Bank, 2003).[5] The indicator accounts for this new form of technological infrastructure, which facilitates the transfer of codified knowledge. The Internet is a new technology that has quickly become the keystone of the Information and Communication Technology, but it was not yet commercially available in 1990. For this reason, we have postponed the beginning of the period to 1994.

2.4. Telephone penetration

Sum of telephone mainlines and mobile phones per 1000 people (ITU, 2001; UNDP, 2001; World Bank, 2003).[6] Telephony, both in fixed and mobile forms, constitutes a fundamental infrastructure for business purposes, and it allows tracing populations with human skills and acquiring technical information.

2.5. Electricity consumption

Number of kilowatts of electricity consumed per hour per capita. This measures the production of power plants and combined heat and power plants, less distribution losses, and own use by heat and power plants (see World Bank, 2003, Table 5.10). This indicator accounts for the oldest technological infrastructure. Electricity consumption is also a proxy measure for the use of machinery and equipment since most of it is generated by electric power.[7]

2.6. Tertiary science and engineering enrolment

Share of tertiary students in science and engineering in the population of that age group (UNESCO, 2002; World Bank, 2003). This variable is a measure of the formation of advanced human capital in science and technology, which represents a necessary requirement for acquiring and managing advanced technological knowledge.

2.7. Mean years of schooling

Average number of years of school completed in the population over 14 (Barro and Lee, 2001; UNDP, 2001; World Bank, 2003). Although this indicator does not consider differences in the quality of schooling across countries, it constitutes a widely used proxy of the level of basic human skill.

2.8. Literacy rate

The percentage of people over 14 who can, with understanding, read and write a short, simple statement on their everyday life (UNDP, 2001; World Bank, 2003). The literacy rate is a necessary precondition for the development of human skills and of basic and advanced human capital.

These eight indicators are those that we will use to analyse the distribution of knowledge across nations and to measure the extent of cross-country differences in the 1990s. The reason for selecting this set of indicators is two-fold. First, from a conceptual point of view, we follow previous studies (Archibugi and Coco, 2004a,b; UNDP, 2001) and argue that these variables represent distinct aspects of the process of national technological accumulation and capacity building, namely the innovative capability of a nation (i.e. patents and scientific articles), its endowment in terms of (new and old types of) technological infrastructures (e.g. Internet, telephony and electricity), and its advanced and basic education levels (i.e. tertiary enrolment, schooling and the literacy rate). Secondly, these indicators are available for a very large number of developed and developing economies in the 1990s. Other indicators could have been chosen instead to measure more specific aspects of national knowledge capabilities, but their country coverage would have been much more limited, due to missing data for several less developed economies.

The advantage of using a relatively large number of indicators is that we are able to provide a more precise characterization of countries' positions in the process of knowledge creation and dissemination than if we were using one single indirect measure such as the TFP. This is particularly important in a large sample that includes countries characterized by very different levels of technological and economic development. It is however necessary to point out two important aspects, and possible limitations, of this set of indicators.

The first is that these variables do not include any measure of international integration. It could be argued that two countries with identical performance in the eight dimensions highlighted above will have different capabilities if their level of international integration is substantially

---

[4] The source of these data is the Science Citation Index generated by the Institute for Scientific Information, which is the most comprehensive and validated available source on scientific publication. It reports information on the scientific and technical articles published in a sample of about 8000 journals selected among the most prestigious in the world. The fields covered are: physics, biology, chemistry, mathematics, clinical medicine, biomedical research, engineering and technology, and earth and space sciences. Article counts are based on fractional assignments; for example, an article with two authors from different countries is counted as one-half article to each country.

[5] Internet users access a worldwide network. They differ from Internet hosts, which are computers with active Internet Protocol (IP) addresses connected to the Internet. The data on users, when available, are preferable to those on hosts for two reasons: first, they give a more precise idea of the diffusion of Internet among the population; second, some hosts do not have a country code identification and in statistics are assumed to be located in the United States, thus causing a bias towards that country.

[6] We have chosen to assign equal weights to fixed and mobile phones because, although they incorporate different degrees of technology, they share the same function.

[7] Other valuable measures of industrial capacity developed, for example, by Lall and his colleagues (see Lall and Albaladejo, 2001; UNIDO, 2002) are available for a smaller number of countries only.

Table 1a
Results of the factor analysis for eight technological indicators, year 1990

	Principal components		Generalized least squares		Maximum likelihood	
	Factor 1: Technological infrastructures and human skills	Factor 2: Creation and diffusion of codified knowledge	Factor 1: Technological infrastructures and human skills	Factor 2: Creation and diffusion of codified knowledge	Factor 1: Technological infrastructures and human skills	Factor 2: Creation and diffusion of codified knowledge
Patents per capita	0.18	0.86	0.23	0.71	0.23	0.71
Scientific articles	0.35	0.86	0.29	0.96	0.29	0.96
Internet penetration ('94)	0.28	0.78	0.32	0.63	0.32	0.63
Telephone penetration	0.86	0.34	0.86	0.34	0.86	0.34
Electricity consumption	0.85	0.29	0.83	0.31	0.83	0.30
Tertiary S&E enrolment	0.74	0.29	0.61	0.40	0.61	0.40
Mean years of schooling	0.80	0.42	0.77	0.44	0.75	0.44
Literacy rate	0.91	0.12	0.87	0.17	0.85	0.17
% of variance explained	46.27	32.00	42.29	30.28	41.77	30.19
Cumulative % explained	46.27	78.28	42.29	72.57	41.77	71.97
χ^2-test[**]	–		26.43 (0.015)[**]		36.01 (0.001)[**]	

Source: Elaboration from Archibugi and Coco, 2004a. *Rotation method*: Varimax with Kaiser normalization.

[**] Significance levels between brackets.

different. One of the key characteristics of technology is precisely its global dimension (Archibugi and Michie, 1995; for a survey, see Carlsson, 2006), and a country well connected and well integrated in the world economy will be able to access and exploit more knowledge and to make better use of it. In the original ArCo database, three indicators of technological integration were, in fact, also considered: inward foreign direct investment, technology licensing payments and import of capital goods (see Archibugi and Coco, 2004a, Table 7). Because of data restrictions, though, it was possible to consider these indicators for a more limited number of countries only, and data for several developing countries were missing. The similarity between the two country rankings was, however, quite evident (linear correlation coefficient between the ArCo index and the ArCo-plus-import-technology index equal to 0.99). This indicates that the advantage of being well integrated into the global economy is also reflected in the values of the eight indicators presented above. In short, by focusing on these eight dimensions, we may lose in theory the effect of global integration of knowledge; the real effect is however rather small and, by restricting our focus to these variables, we are able to consider a much larger set of economies, including also many developing countries.

A second major difficulty in using these indicators is that they represent to a great extent strictly related (and not easily separable) aspects of countries' technological capabilities. Variables measuring the creation and diffusion of new knowledge, technological infrastructures and human skills tend in fact to be highly correlated, as they all constitute complementary dimensions of national technological capabilities.[8] Thus, before proceeding to the core of our

empirical analysis, the study of cross-country differences in technology, it is important to reduce this large set of indicators to a smaller number of *distinct* (not overlapping) dimensions.

A factor analysis has been performed in order to identify the explanatory variables that are better suited to analyse the cross-country distribution of knowledge, i.e. those variables that better discriminate between countries' technological levels.[9] The purpose of the factor analysis is to extract a smaller number of factors that are able to account for most of the variance in the sample.[10]

Tables 1a and 1b present the results of the factor analysis for our ArCo dataset for the years 1990 and 2000, respectively. Three different methods of factor extraction have been used: (i) principal components; (ii) generalized least squares; and (iii) maximum likelihood.[11] All three methods identify two major factors, which appear to be stable over time between 1990 (Table 1a) and 2000 (Table 1b). These two factors taken together explain a large percent-

---

[8] Relatedly, some of these indicators are also positively correlated with GDP per capita, given that technological infrastructures, human skills and innovative capabilities are closely linked to the overall level of economic development of a country. This is a more general limitation of the set of indicators that we use in this paper. This limitation is, however, not easy to overcome, given that most indicators of technological capabilities that are available for a large sample of countries do also reflect other aspects of

the development process like consumption and investment patterns, as for instance is the case for our indicators of telephony, electricity consumption and education levels.

[9] Interestingly, in a comprehensive survey of growth theory, Temple (1999, p. 148) argues: "Despite some interest from development economists, simple techniques for data reduction like factor analysis and principal components have been largely ignored by recent growth researchers. Their use seems to have a great deal of potential." For previous papers applying factor analysis to the study of economic growth and development, see Adelman and Morris (1965) and Temple and Johnson (1998).

[10] The eight indicators of countries' technological capabilities have been standardized before entering the factor analysis. The general formula used to standardize the indicators adopts the "distance from the best and the worst performers" method of standardization, which is the same used for the indicators composing the Technology Achievement Index (UNDP, 2001; Desai et al., 2002) and the ArCo Technology Index (Archibugi and Coco, 2004a,b). Different methods of standardization could have been used, but the choice of a different standardization method would have not affected the results of the factor analysis.

[11] The use of these three different methods of factor extraction enables to assess the robustness of the results, as different algorithms may sometimes lead to extracting different factors.

Table 1b
Results of the factor analysis for eight technological indicators, year 2000

	Principal components		Generalized least squares		Maximum likelihood	
	Factor 1: Technological infrastructures and human skills	Factor 2: Creation and diffusion of codified knowledge	Factor 1: Technological infrastructures and human skills	Factor 2: Creation and diffusion of codified knowledge	Factor 1: Technological infrastructures and human skills	Factor 2: Creation and diffusion of codified knowledge
Patents per capita	0.17	0.89	0.22	0.79	0.22	0.78
Scientific articles	0.34	0.89	0.30	0.95	0.31	0.95
Internet penetration	0.38	0.81	0.40	0.74	0.41	0.74
Telephone penetration	0.86	0.34	0.86	0.35	0.87	0.34
Electricity consumption	0.84	0.34	0.82	0.34	0.83	0.34
Tertiary S&E enrolment	0.65	0.51	0.59	0.53	0.60	0.53
Mean years of schooling	0.79	0.45	0.75	0.47	0.74	0.47
Literacy rate	0.91	0.10	0.84	0.17	0.82	0.17
% of variance explained	45.12	37.21	41.83	35.59	41.55	35.32
Cumulative % explained	45.12	82.33	41.83	77.42	41.55	76.87
χ^2-test[**]		–	21.04 (0.072)[**]		27.57 (0.010)[**]	

Source: As for Table 1a. *Rotation method*: Varimax with Kaiser normalization.

[**] Significance levels between brackets.

age of the variance in the sample (between 72% and 82%), and the goodness of fit of the two-factor solution is confirmed by the χ^2-test (reported in the last row of the tables for the GLS and ML methods). This is significant at conventional levels, thus showing that these two factors account for most of the variability in the data.[12] Our interpretation of the factor solutions is the following.

Factor 1 may be interpreted as a broad measure of *technological infrastructures and human skills*. This principal component in fact loads very high on all the variables measuring technological infrastructures (telephone penetration and electricity consumption) and those measuring education levels and human skills (tertiary science and engineering enrolment, mean years of schooling, and literacy rate). The extraction of this first factor shows that infrastructures and human skills are highly correlated and strongly complementary to each other. Both of them contribute to define each country's *absorptive capacity*, i.e. its ability to imitate and to implement foreign advanced technologies, which is a key requirement for developing economies that try to reduce the huge gap that separates them from the technological frontier (Abramovitz, 1986). Factor 1, a linear combination of these five variables, accounts for around 40–45% of the variance in the sample in both periods, and it therefore appears as a very relevant dimension to investigate cross-country differences in technological capabilities in the 1990s.

Factor 2 may be interpreted as a measure of *the creation and diffusion of codified knowledge*. It is highly correlated with variables measuring the creation of codified techno-

logical knowledge by private firms and the public sector (i.e. patents per capita and the production of scientific articles), and its diffusion (measured by the Internet penetration). The factor indicates that the creation and diffusion of advanced technologies are to a large extent related and complementary aspects of countries' technological activities. Both contribute to determine the *innovative capability* of nations, which is a fundamental requirement to compete in the modern knowledge-based economy, particularly for countries that are already close to the technological frontier. Factor 2 is a linear combination of these three variables, and it accounts for about 30–37% of the variance in the sample. Thus, together with factor 1, it appears to be a very important aspect to study differences in technological capabilities in the world economy.

Fig. 1 reports the kernel density estimates of the statistical distributions of these two principal components in 1990 and 2000. The graphs show clearly that the shapes of the distributions of the two factors are very different. The thick and long left tail of factor 1 suggests the existence of a great cross-country variability of technological infrastructures and human skills within the large group of less developed economies. By contrast, factor 2 shows a long right tail, which indicates that innovative capabilities are very low for most developing countries in the sample, while they vary significantly within the group of middle income and advanced countries. Given that the purpose of this paper is to analyse differences in technological capabilities in a very large sample of 131 countries, including both developing and developed economies, both factors identified by the principal component analysis will be used in the subsequent empirical analysis.

3. The technology clubs and their dynamics over the 1990s

We will now use these two principal components to investigate the existence and the characteristics of different technology clubs in the world economy in the 1990s. This section presents the results of a cluster analysis that divides

---

[12] The same tests for the three-, four-, and five-factor solutions yield no significant results. This suggests that the extraction of additional factors after the second contributes only marginally to the overall percentage of variance explained, and the two-factor solutions may be regarded as the most efficient, in the sense of explaining a large portion of the variability in the sample by using only a small number of factors. This conclusion is confirmed by graphical inspection of the Scree plot, which also supports the choice of the first two factors as the most relevant for our subsequent analysis.
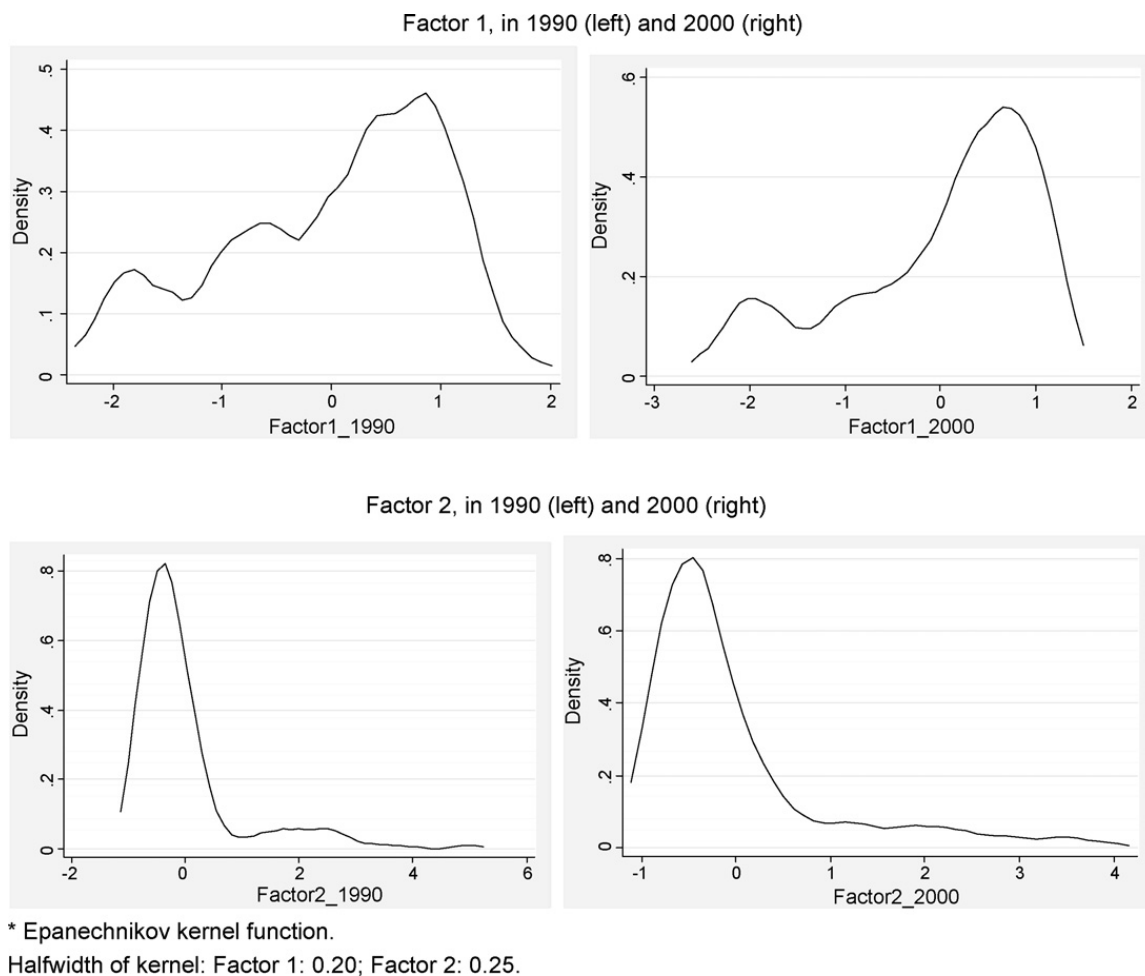
Fig. 1. Kernel density estimates of the two principal components (Epanechnikov kernel function). Halfwidth of kernel: factor 1: 0.20; factor 2: 0.25.

the world economy into a few country groups characterized by different levels of knowledge stocks and technological development.

Two different specifications of the cluster analysis have been tested. In the first, the two principal components extracted by the factor analysis in the previous section have been used as inputs in the clustering algorithm. In the second, two variables have been used instead: the literacy rate (corresponding to factor 1) and the number of scientific articles per capita (corresponding to factor 2). These two variables have been selected because (i) they are more strongly correlated with each component of the factor analysis (see Tables 1a and 1b), and (ii) their distribution shapes are more similar to those of the two components. In other words, this means that these two indicators are those that best represent (and most closely correspond to) the factors extracted in the factor analysis.[13] In fact, the literacy rate variable varies substantially within the large group of developing economies (as the first principal component), while the scientific articles indicator has a great deal of cross-country variation when we consider middle-income and more advanced countries (as the sec-

ond component). Both aspects are therefore relevant in our large cross-country sample.

As is well known, the results of cluster analysis may change substantially when different methods are applied to the same dataset. In order to overcome this problem, we have run around thirty different clustering algorithms. For each of the two specifications, we have proceeded in three steps. First, three distinct *hierarchical agglomerative* methods have been applied (between groups linkage, within groups linkage, Ward's method), and three different ways to measure the distance between cases have been used (Euclidean, squared Euclidean, and cosine distance). Secondly, after having identified the main resulting clusters, an *iterative partitioning* method (K-means procedure) has been applied to check the robustness of the results of the hierarchical agglomerative analysis. Additionally, each clustering algorithm has been run by using as inputs the factors (and the corresponding variables) in both 1990 and in 2000, so as to assess the stability of the resulting clusters over time.

The methodology that we have employed has three advantages: (i) it makes it possible to check the robustness of the results to changes in the variables used as inputs, in the clustering method and in the distance measurement adopted; (ii) it allows investigating whether the resulting clusters are stable over time between 1990 and 2000, in terms of both cluster characteristics and cluster member-

---

[13] Both variables have been standardized before entering the clustering algorithms, as is customary in cluster analysis.

Table 2
Main characteristics of the three technology clubs

	Cluster 1: Advanced		Cluster 2: Followers		Cluster 3: Marginalized	
	1990	2000	1990	2000	1990	2000
Patents granted in USPTO[a]	69.45	97.37	4.29	6.81	0.02	0.03
Scientific articles[a]	627.36	670.65	68.56	90.54	4.94	5.63
Internet users (1994 and 2000)[b]	26.67	289.77	2.48	57.32	0.01	3.51
Fixed and mobile telephones[b]	516.78	1,055.92	163.07	404.72	13.36	47.14
Electricity consumption (kWh per capita)	9411.5	10,450.9	2584.1	2989.4	265.8	318.5
Tertiary S&E enrolment ratio	10.87	17.31	6.68	9.33	1.28	2.06
Mean years of schooling[c]	9.91	10.44	6.56	7.06	3.42	3.93
Literacy rate[c]	98.66	98.80	91.29	93.86	58.01	67.57

Source: As for Table 1a. The list of countries included in each cluster is reported in Appendix A.

[a] Per million people.
[b] Per thousand people.
[c] Population over 14.

ship; (iii) the hierarchical agglomerative method employed is able to identify *endogenously* the number of clusters that forms the best partition of the dataset.

All of the clustering algorithms have produced *three* main clusters, and this result appears to be robust to modifications in the method applied, in the definition of distance measurement adopted, and in the variables used as inputs in the analysis. The existence of three different groups of countries is also apparent from a graphical inspection of the kernel density estimates of the two principal components, which indicate that both of them, and particularly factor 1, are characterized by a three peaks distribution (see Fig. 1, Section 2).

The three resulting clusters are to a large extent the same in 1990 and 2000, with only a few changes of cluster memberships between the two periods. The major characteristics of these three clubs are presented in Table 2, and the cluster membership is indicated in Appendix A. Table 3 reports, in addition, the technology gap between the clubs and its rate of change over the 1990s. The three groups are briefly described as follows.

3.1. Cluster 1: advanced (high infrastructures and skills, high innovation)

This is the group of more technologically advanced countries, composed of a small set of industrialized economies (between 15 and 21 countries, less than 15% of the sample's population) that hold approximately 40%

of the world GDP. It comprises the traditional leaders, US and Japan, Continental and Northern European economies, and Western offshoots (Australia, Canada, Israel and New Zealand). The cluster membership is quite stable over time, the main difference being the entry of a few very dynamic Asian NICs (Hong Kong, South Korea and Singapore) into this restricted club in 2000.

Table 2 shows that at the beginning of the period the group was characterized by high levels of creation and diffusion of codified knowledge (on average, around 69 patents, 627 scientific articles and 27 Internet users per capita), well-developed technological infrastructures (in terms of both telephones and electricity consumption), and high levels of basic and advanced education (nearly 11% tertiary enrolment ratio in science and engineering, 10 mean years of schooling, and 99% literacy rate).

The cluster does also prove to be very dynamic over time, with rapid growth rates of technological capabilities between 1990 and 2000, particularly in terms of patents (+40%), Internet users (+986%), telephones (+104%) and tertiary enrolment ratios in science and engineering (nearly +60%). The technological frontier has thus advanced rapidly in the last decade.

3.2. Cluster 2: followers (medium-high infrastructures and skills, low innovation)

This is a larger group, composed of around 70 countries (around 27% of the sample's population in 1990) that

Table 3
The technology gap between the three clubs, and its change over the 1990s

	Advanced vs. Followers			Followers vs. Marginalized		
	1990	2000	%Change	1990	2000	%Change
Patents	16.18	14.29	−11.6	190.3	237.4	+24.7
Scientific articles	9.15	7.40	−19.0	13.87	16.09	+16.0
Internet users (1994 and 2000)	10.77	5.05	−53.0	270.6	16.32	−94.0
Fixed and mobile telephones	3.17	2.61	−17.7	12.20	8.58	−29.6
Electricity consumption	3.64	3.50	−4.0	9.72	9.39	−3.4
Tertiary S&E enrolment ratio	1.63	1.85	+13.8	5.21	4.54	−12.9
Mean years of schooling	1.51	1.48	−2.1	1.92	1.79	−6.4
Literacy rate	1.08	1.05	−2.6	1.57	1.39	−11.7

Source: As for Table 1a. The first and second columns report the ratio between technological capabilities in the advanced and followers clusters in 1990 and 2000, and the third indicates the rate of change of the technology gap in the period. Similarly, the fourth and fifth columns report the ratio between technological capabilities in the followers and marginalized clusters in 1990 and 2000, and the sixth shows the rate of change of the technology gap over the 1990s.

accounted for 36% of the world GDP at the beginning of the period. The cluster membership is quite stable between 1990 and 2000, and its core is constituted by catching-up economies from South-East Asia, Southern Europe, the Middle East, Latin America, plus the large group of Former Socialist countries.

There are only a few notable changes to this cluster composition over time. In 2000, the main exit consists of a small set of dynamic Asian and Continental European economies that abandon the group and move towards the advanced club, while the entry is that of a small number of fast catching-up developing countries, from Asia, the Middle East, Central America and Africa, which rapidly improved their technological capabilities in the 1990s.

Admittedly, this second club represents a rather heterogeneous group of economies. Compared to the advanced cluster, the group shows on average a much lower ability to create and to imitate advanced knowledge, as measured by the number of patents, scientific articles and Internet users. In fact, Table 3 indicates that at the beginning of the 1990s the innovation gap between the advanced and the followers group is huge (around 16:1 for patents, 9:1 for articles and 11:1 for Internet users). Over the decade, the knowledge distance has gradually decreased in terms of patents and articles, and much more rapidly in terms of Internet users. However, the innovation gap *vis-à-vis* the economies in the advanced cluster does remain considerable at the end of the decade. On the other hand, the technological distance between this second group and the more advanced one is significantly lower in terms of infrastructures (around 3:1 for telephones and electricity), and education levels (between 1.5:1 and 2:1), in both 1990 and 2000.

3.3. Cluster 3: marginalized (low infrastructures and skills, low innovation)

This is the largest group of countries, accounting for more than 60% of the sample's population in 1990, but producing only around 23% of the world GDP at the beginning of the decade. The core of this cluster consists of large Asian economies plus nearly all the African countries. The group membership is rather stable in the 1990s, although a restricted number of dynamic economies have managed to catch up and to join the more advanced followers cluster in 2000 (e.g. China, Indonesia, Vietnam, Iran, Oman, and a very few countries from Central America and Africa).

In terms of the ability to create advanced knowledge and to imitate new technologies, this group is not only remarkably far from the advanced club, but also quite distant from the follower countries in the second cluster. Table 3 shows in fact that at the beginning of the period the distance between the followers and the marginalized groups is around 190:1 in terms of patents, 14:1 for scientific articles, and 270:1 for Internet users. During the decade, the gap diminished very rapidly in terms of Internet users, but it significantly widened for patents and articles.

Regarding the indicators of technological infrastructures and human skills, the distance *vis-à-vis* the followers cluster at the beginning of the period is equally striking, particularly in terms of telephones (12.2:1), electricity consumption (9.7:1) and tertiary enrolment ratio in science and engineering (5.2:1). Over the 1990s, the gap in terms of technological infrastructures and human skills decreased slowly, but it still remained huge at the end of the decade.

In short, the results of the cluster analysis lead to the identification of three technology clubs with markedly different levels of technological development. The characteristics of these technology clubs closely resemble those of the 'innovation', 'imitation' and 'stagnation' groups identified by the recent model of Howitt and Mayer-Foulkes (2005). Our findings indicate the existence of two large knowledge gaps in the world economy: the first refers to the great distance that separates the group of followers from the technological frontier, particularly in terms of innovative capabilities; the second refers to the impressive gap that separates the marginalized from the followers clubs, in terms both of innovative capabilities and of infrastructures and human capital.

The three clubs do not only differ in terms of their levels of technological development, but also with respect to the dynamics that they experienced over the 1990s. By looking at the third column in Table 3, we observe in fact that the club of followers on average decreased its distance from the advanced group for nearly all the indicators. This catching-up process was particularly rapid not only in terms of Internet users and telephony (reflecting the worldwide diffusion of ICTs), but also with respect to innovative capabilities (patents and scientific articles), which is precisely the aspect where the technology gap between the followers and the advanced clubs is more evident. The only exception to this general trend of convergence refers to the tertiary enrolment ratio, whose gap *vis-à-vis* the frontier increased by around 14%, due to the very rapid increase of tertiary education in most advanced countries in the period.

However, when we turn the attention to the dynamics of the marginalized *vis-à-vis* the followers club (see the sixth column of Table 3) the picture is quite different. Here, convergence is rapid only in terms of Internet and telephony, and rather slow for all other indicators of technological infrastructures and education levels (i.e. electricity consumption, tertiary enrolment ratio, years of schooling and literacy rate). On the other hand, the innovation gap between the marginalized and the followers clubs significantly increased (patents: +24.7%; scientific articles: +16%), indicating that the group of marginalized economies was not able to improve its innovative capabilities, while the other two groups have been very dynamic in this respect.

This result is a reason for concern for the future of the world economy. Given that innovation is a key requirement to compete in the modern knowledge-based economy, the increasing polarization between rich and poor countries in terms of innovative capabilities may in fact lead to greater disparities in income and GDP per capita in the years ahead.[14]

---

[14] The present paper does not aim at analysing in further detail the relationships between absorptive capacity, innovative capability and economic performance (Pianta, 1995; Castellacci, in press). However, a brief

4. Econometric test of the technology clubs classification

The results presented in the previous section, on the existence and the dynamics of these country clubs in the 1990s, have been obtained through a cluster analysis, and their validity depends therefore on the statistical precision of the latter. This section presents the results of an econometric test that aims at assessing the statistical precision of our club classification.

This test is relevant for one major reason. In applied growth theory, this type of clustering exercise to identify different country groups has recently become very popular. In the last few years, several empirical studies have in fact proposed different classifications of countries into convergence clubs (Durlauf and Johnson, 1995; Desdoigts, 1999; Hobijn and Franses, 2000; Johnson and Takeyama, 2001; Papageorgiou, 2002; Fiaschi and Lavezzi, 2003; Canova, 2004). However, the various exercises differ substantially in terms of (i) the variables used as discriminatory factors, (ii) the sample considered (time span and number of countries included), (iii) the clustering method employed, and consequently (iv) the results (i.e. the number and composition of country clubs). This leads to a couple of relevant questions. In general terms, how can we compare the statistical precision of different classification exercises? And more specifically, is the precision of the club classification presented in this paper comparable to that of previous related exercises? Given the recent flourishing of applied studies in the convergence clubs literature, we believe it is necessary to find some way to measure and assess the statistical precision of different classifications, so as to be able to evaluate their explanatory/classificatory power and reflect upon the validity of the underlying theoretical interpretation.

Hansen (2000) has recently proposed a method to deal with this problem.[15] The method computes confidence intervals of threshold parameters for the class of so-called threshold regression econometric models. That is, given the results of a clustering exercise obtained through, say, a regression tree analysis (e.g. Durlauf and Johnson, 1995), Hansen's method makes it possible to calculate the statistical precision of the resulting partition of the sample. However, the main problem with this state-of-the-art approach is that both the clustering method and the test for calculating the precision of the results are based on a cross-country growth regression model (typically a standard augmented Solow model). If the latter is misspecified, for instance by neglecting some relevant technology-related factors, then the thresholds that partition the sample would

not necessarily be those that lead to the best partition of the sample space.[16]

We propose here a different approach, which is not based on a cross-country growth regression model, and therefore does not impose any a priori structure on the data. The method is based on the estimation of a multinomial logit model, which estimates the strength of the association between the discriminatory factors (input variables) and the resulting country classification. This makes it possible to evaluate the statistical fit of different classification exercises, regardless of the sample, variables and method used to obtain them. In the remainder of this section, we employ this methodology to assess the statistical precision of our club classification. Appendix B then uses the same method to analyse the results of Durlauf and Johnson's (1995) classification exercise, which is the best-known study in the convergence clubs literature.

The test for our club classification is constructed as follows. The dependent variable is the categorical variable 'technology clubs', which takes value 2 for 'advanced' countries, 1 for 'followers', and 0 for 'marginalized' economies. The purpose is to estimate the relationship between the choice of assigning country i to cluster j (where j = 0, 1, or 2) and a set of explanatory variables, which are those used as inputs in the cluster analysis. An OLS approach cannot be used in this case, as the explanatory variables are continuous, while the dependent is a categorical variable that takes only three values. The standard way to solve this problem is to estimate a multinomial logit (MNL) model (Scott Long, 1997; Peracchi, 2001). This is commonly expressed as:

$$\Pr\left\{Y_i = j\right\} = \frac{\exp(\beta_j^{\mathrm{T}} X_i)}{1 + \sum_k \exp(\beta_k^{\mathrm{T}} X_i)} \quad \text{for } j = 2, 3, ..., J \quad (1)$$

$$\Pr\left\{Y_i = 1\right\} = \frac{1}{1 + \sum_k \exp(\beta_k^{\mathrm{T}} X_i)} \quad \text{for } j = 1 \quad (2)$$

where X_i is a vector of characteristics specific to country i, and β_j is a vector of coefficients specific to cluster j.[17] The country-specific characteristics used as explanatory variables in the model are those previously identified by the factor analysis, and then used as inputs in the cluster analysis, i.e. factor 1 (technological infrastructure and human skills) and factor 2 (creation and diffusion of codified knowledge).[18] The results of the estimations are presented in Table 4a (for the year 1990) and Table 4b (for the year 2000).[19]

---

look at the coefficients of correlation between our set of indicators and the GDP per capita in the period corroborates the above statement about the crucial importance of innovation. Both factors resulting from the principal component analysis are in fact positively correlated with the GDP per capita. However, the first factor, representing technological infrastructures and human skills, has a coefficient of correlation that is lower than the corresponding coefficient for the second factor, which represents instead innovative capabilities (in 1990: 0.51 for factor 1 and 0.65 for factor 2; in 2000: 0.57 for factor 1 and 0.64 for factor 2). The relationship between different components of the process of technological accumulation and economic performance is indeed a relevant topic for further research.

[15] See also the related work of Los (2006).

---

[16] For instance, Masanjala and Papageorgiou (2004) have shown that, by simply changing the functional form of the Cobb-Douglas specification of the growth equation, the resulting classification of countries into convergence clubs differ substantially from the one originally proposed by Durlauf and Johnson (1995).

[17] Eqs. (1) and (2) are non-linear, and require an iterative solution. This is based on the method of maximum likelihood, whose maximum is typically found by Newton's method in just a few iterations.

[18] We have additionally estimated a MNL model by using as explanatory variables the number of scientific articles and the literacy rate, because these are the two variables that have been used as inputs in the second specification of the cluster analysis (see Section 3). The results of these additional estimations are largely consistent with those presented here, and are available on request.

[19] The models presented in Tables 4a and 4b, as well as the models presented in Table 5 below, do not include an intercept term. The reason is

Table 4a
Results of the MNL test for the 'technology clubs' classification, year 1990 (Wald statistics between parentheses)

	Cluster 2: Followers	Cluster 3: Marginalized	Likelihood ratio test
Factor 1: Technological infrastructures and human skills	24.47 (133.44)[***]	−6.93 (98.27)[***]	185.35[***]
Factor 2: Creation and diffusion of codified knowledge	−16.68 (140.96)[***]	−1.50 (16.03)[***]	98.32[***]
Pseudo R^2	Cox and Snell: 0.878	Nagelkerke: 0.988	R_{LA}: 0.944
Percentage of correctly predicted cases	Cluster 2: 98.5%	Cluster 3: 100%	Overall: 99.1%

Dependent variable "Technology clubs": $Y = j$, where $j = 2$ for 'advanced', $j = 1$ for 'followers', and $j = 0$ for 'marginalized' countries.

[***] Significance at the 0.01 level.

The multinomial logit model is essentially a "linked set of binary logits" (Scott Long, 1997). In our case, the model simultaneously estimates two binary logits, i.e. a vector of coefficients β_j for the followers and the marginalized groups of countries relative to the advanced cluster, which has been used as the reference category (the reason why the advanced cluster is not reported in a separate column like the other two groups). Note that the choice of the baseline category does not affect the results in any way, so that any other cluster could have been chosen instead.

Looking at the upper part of Tables 4a and 4b, we notice that both factor 1 and factor 2 turn out to be significant in the model, as confirmed by the likelihood ratio test for each single regressor. This is a χ^2-distributed statistic that compares the final model with a reduced model, omitting each single explanatory variable. However, the relative strength of the explanatory variables differs slightly in each cluster. In a MNL model each estimated coefficient measures the proportional change in the log of the odds-ratio of the dependent variable when the kth regressor changes by one unit. In other words, if the estimated coefficient β_k is positive (negative), the likelihood of that response category will increase (decrease) by a factor of β_k for any unit change of the kth regressor.

The estimated coefficients for the models in both Tables 4a and 4b confirm the basic characteristics of the clubs. Factor 2, a general measure of the creation and diffusion of codified knowledge, is negative and significant in both clusters, meaning that a greater ability to innovate and to imitate lowers the likelihood that a country is classified as followers or marginalized (i.e. it increases the probability that it is classified as advanced, which is the reference category).

The magnitude of this estimated coefficient is much stronger for the followers than for the marginalized group, thus confirming that factor 2 is a more relevant variable to discriminate between rich and middle-income countries. Furthermore, factor 2 appears to have a stronger effect in 2000 than in 1990, thus suggesting that the creation and diffusion of codified knowledge is gradually becoming a more crucial discriminatory factor in the modern knowledge-based world economy. During the 1990s, in fact, only those follower countries that have been rapidly able to enhance their capability to innovate and to imitate, such

as Asian NICs, have been able to shift from the followers to the advanced club.

When the attention is turned to the other explanatory variable, factor 1 (a broad measure of technological infrastructures and human skills) turns out to be positive and significant for the followers club in both 1990 (Table 4a) and 2000 (Table 4b), meaning that greater levels of infrastructures and human capital increase the probability that a country is classified as follower rather than as advanced. The reason may be that the restricted group of advanced economies have already reached very high levels of infrastructures and human skills, while follower countries, although quite close to the knowledge frontier, have still some scope for improvement in this respect. So, an increase in factor 1 is more relevant for the followers than for the advanced cluster, thus increasing the probability that a country is assigned to the former rather than to the latter group.

The same does not hold for marginalized economies, however, because their distance from the technological frontier in terms of infrastructures and human skills is too wide. In this case, factor 1 turns out to be negative and significant in the estimations in both 1990 and 2000. This means that the higher the level of infrastructures and human skills, the lower is the likelihood that a country is classified as marginalized, and conversely the higher the probability that it is classified as advanced (the baseline category).[20]

The lower part of Tables 4a and 4b reports indexes of the overall model fitting, such as the Cox–Snell, the Nagelkerke and R_{LA} pseudo R-squared.[21] All three indexes show a very good overall fit of the model, thus confirming the precision of the results of the cluster analysis for both 1990 and 2000. This may be interpreted as saying that the explana-

---

that if both explanatory variables take simultaneously value 0, say in the case of an extremely poor country, then the dependent variable will be 0 as well (i.e. the country belongs to the marginalized cluster).

[20] It is worth observing that the explanatory variables that we have included in the econometric model are not intercorrelated, so that no multicollinearity problem arises in the estimations. The reason is that the explanatory variables in the MNL model are the principal components extracted by the factor analysis (see Section 3), which are by construction orthogonal to each other. Thus, they represent two distinct (not overlapping) dimensions of countries' technological capabilities.

[21] The Cox–Snell pseudo R-squared is an analogous to the interpretation of multiple R-squared based on the likelihood, but its maximum can be less than 1. The Nagelkerke pseudo R-squared is a modification of the former index so as to assure that it can vary between 0 and 1. Finally, the R_{LA} index is a pseudo R-squared adjusted to take into account the number of explanatory variables in the model (similarly to the adjusted R-squared in OLS regressions). Note that all these indexes are *pseudo R-squareds*: they aim at measuring the strength of association in the regression model, but cannot be interpreted as measuring the percentage of variance explained by the model, as is the case for the R-squared in OLS regressions.

Table 4b
Results of the MNL test for the 'technology clubs' classification, year 2000 (Wald statistics between parentheses)

	Cluster 2: Followers	Cluster 3: Marginalized	Likelihood ratio test
Factor 1: Technological infrastructures and human skills	27.82 (95.21)[***]	−3.94 (101.44)[***]	121.52[***]
Factor 2: Creation and diffusion of codified knowledge	−51.12 (108.60)[***]	−2.09 (17.04)[***]	158.54[***]
Pseudo R^2	Cox and Snell: 0.881	Nagelkerke: 0.991	R_{LA}: 0.954
Percentage of correctly predicted cases	Cluster 2: 98.7%	Cluster 3: 100%	Overall: 99.2%

Dependent variable "Technology clubs": $Y = j$, where $j = 2$ for 'advanced', $j = 1$ for 'followers', and $j = 0$ for 'marginalized' countries.
[***] Significance at the 0.01 level.

tory variables are strongly associated with the dependent variable, i.e. the probability that country i is assigned to cluster j. This is a first way to evaluate the validity of the results of the cluster analysis: the higher the pseudo R-squared indexes, the stronger the relationships between explanatory and dependent variables, the more precise are the results of the cluster analysis.

An additional way to measure the precision of the classificatory exercise presented in the previous section is to look at the correctly predicted percentage of cases, reported in the lower part of the tables. These provide a useful summary of the predictive ability of the multinomial logit model.[22] The classification table shows that the taxonomy proposed in the previous section has on the whole a very high percentage of correctly predicted cases (around 99%).

On the whole, the estimation of the multinomial logit model supports the validity of the results of the cluster analysis presented in the previous section, and thus the existence of three distinct technology clubs in the world economy in the 1990s. The significance of the single estimated coefficients, the overall strength of association between dependent and explanatory variables, and the predictive ability of the model as measured by the classification table, all confirm that the cluster analysis has identified three groups of countries which significantly differ between themselves with respect to the chosen explanatory variables.

For benchmark purposes, it is interesting to use the same MNL methodology to assess the statistical precision of the best-known previous study in this field, namely Durlauf and Johnson's (1995) work on convergence clubs. The results of this additional test are presented in Appendix B. In a nutshell, these results indicate that the classificatory power of the new knowledge clubs partition proposed in this paper is comparable to that of Durlauf and Johnson's well-known classification. We take this as an indication of the validity and the statistical precision of the results presented in Section 3, on the existence and the dynamics of the three country clubs in the 1990s.

While providing encouraging evidence in support of our study, these econometric results also suggest a more general reflection. Economic and technological factors are both relevant aspects in the analysis of cross-country differences. The convergence clubs literature has traditionally focused on the former, while the Schumpeterian approach adopted in this paper emphasizes the latter. From somewhat different perspectives, both of them lead to comparable empirical results in terms of country grouping and classifications. The implication of this is not to argue that one approach is better than the other, but rather to point out that both of them can provide useful insights in the study of the development process. Economic growth is driven by the co-evolution of technological and economic factors, and mainstream growth theory should therefore re-consider the role of technological knowledge and its complex interactions with economic factors.

5. Conclusions and policy implications

The paper has carried out an empirical investigation of cross-country differences in the ability to exploit the process of knowledge creation and dissemination in a large sample of developed and developing economies in the 1990s. The results show the existence of three knowledge clubs characterized by markedly different levels of technological development. The technological distance between the clubs is huge with respect to all the knowledge-related aspects considered in the study, namely innovative capability and the levels of technological infrastructures and human skills. Over the decade, these country groups have experienced different dynamics of technological change. The followers club has in fact come closer to the technological frontier, while the marginalized group has experienced an enlargement of its gap in terms of innovative capabilities *vis-à-vis* the rest of the world.

What is the theoretical interpretation of these empirical results, and what are the implications? Four aspects deserve to be emphasized. First, from the point of view of the convergence clubs literature in applied growth theory, our findings support the hypothesis that the process of knowledge creation and dissemination, and countries' different ability to take advantage of that, is a major factor determining the existence of convergence clubs. While all recent studies in this field have focused on a variety of economic factors that may generate country clubs, the Schumpeterian approach followed in this paper shows that countries differ, first and foremost, in terms of their ability to innovate and to imitate advanced technologies. The focus on knowledge creation and transfer capabilities leads to a precise three-club partition of the world economy, whose classificatory power is comparable to that of the well-

---

[22] The classification table is obtained as follows. For each country i and cluster j, the model estimates $Pr Y_i = j$, i.e. the likelihood that country i is assigned to cluster j given the set of country-specific characteristics X_i. If this estimated probability is higher than a threshold value (0.5 by default), then country i is assigned to cluster j. This set of assignments is then compared to the values of the dependent variable, that is the actual cluster membership of our technology clubs. Finally, the predicted and the actual cluster membership are compared, and the percentage of correctly predicted cases for each cluster is computed.

known study by Durlauf and Johnson (1995). Future studies on the convergence clubs hypothesis, instead of exclusively focusing on standard economic factors, should therefore more carefully consider the crucial role of knowledge and its complex interactions with the process of capital accumulation.

Secondly, from the point of view of the innovation literature, our results bring new empirical evidence on the existence of three distinct groups of national systems of innovation. Admittedly, our indicators provide only a stylised measure of national systems, and do not take into due consideration the great variety of dimensions that characterize the creation and diffusion of new knowledge within each national economy.[23] Nevertheless, the empirical result indicating that there are three markedly different clubs of national systems that follow quite distinct technological trajectories is an interesting stylised fact that needs to be explained by the innovation systems literature. Empirical studies of national systems have in fact so far almost exclusively focused on a restricted group of rich countries (our advanced club), and frequently neglected the developing world. The extension of the innovation systems literature to the study of innovative patterns and performance in less developed economies constitutes an important challenge for the field of innovation studies.

Thirdly, it emerges quite clearly that the differences across countries are well captured by two factors, as indicated by the results of the factor and cluster analyses. We have labelled the first technological infrastructures and human skills, and the second creation and diffusion of codified knowledge. The results highlighted by the cluster analysis are entirely consistent with the existing literature on technological change and, in fact, seem to reconcile the views held by various approaches. When development economics is concerned with knowledge transfer, it naturally tends to stress the importance of absorptive capacity, which in turn requires infrastructures and human skills (Bell and Pavitt, 1997). The Schumpeterian tradition, on the contrary, is more likely to focus on innovative capabilities, reflected in the creation and diffusion of advanced codified knowledge. The empirical analysis has shown that, when we consider a very large sample of countries in the world economy, both factors are relevant to identify the position of nations, and that each of them has a specific menu.

Fourthly, when we consider the dynamics experienced by the three clubs over the decade, our results cannot easily be interpreted to support one theoretical paradigm or another, and should thus be taken with caution. The time span considered in this paper, the 1990s, is in fact too short a period to provide any robust generalization on the convergence or divergence behaviour of different country groups. The results suggest however some interesting indications on the relative dynamics of the three clubs. The followers group has come closer to the technological frontier in terms of most of the technology variables, and this is certainly an encouraging indication for several middle-income

countries that are trying to catch up with the industrialized world. By contrast, the large club of marginalized economies has experienced an increase in its gap in terms of innovative capability, and this is a reason of concern for less developed countries. In a Schumpeterian perspective, innovative capabilities constitute a key engine of growth, and an important source of cumulativeness in the dynamics of economic systems. A greater innovation gap today is likely to lead to a greater income gap tomorrow. The interaction between technological and economic factors would thus possibly drive the countries further and further apart, and such a Myrdalian cumulative causation could therefore lead to greater disparities in the years ahead.

However, as previously said, the short time span considered in this paper does not make it possible to forecast with certainty such a gloomy picture for the developing world, and other elements suggest taking a more cautious and more optimistic point of view. Human capital and technological infrastructures are gradually improving, and ICT-related technologies are diffusing rapidly. Hopefully, what we are observing today could therefore be interpreted in terms of a Kuznets curve, where inequality is widening in this phase of rapid technological change but will later narrow when the international diffusion of advanced technologies will eventually benefit less developed countries as well.

Two main policy implications can be drawn from our empirical results. The first concerns the uneven distribution of knowledge. Not only does each country have a specific combination of knowledge resources, but the various ingredients of knowledge creation and transfer capabilities appear as largely complementary rather than substitutes. The combination is strictly associated with the overall level of technological and, more broadly, economic and societal development. The combination of knowledge resources, and not just their level, does play an important role in positioning each national system of innovation. Any catching-up country should not just observe and try to imitate those who are at the top of the league, but rather trace its own itinerary, taking into account the existing knowledge capabilities and the possibility of successfully upgrading its position. Specifically, the building up of the absorptive capacity of a nation through the improvement of its basic education levels and technological infrastructures should be a necessary prerequisite for the subsequent development of innovative capabilities and, hence, of a successful and sustainable economic performance.

The second policy implication refers to the overall evolution of national systems. How many countries manage to upgrade their competencies and to 'jump' into the higher club? One significant fact is that no country slipped back into a lower group. By contrast, some countries managed to exit the group to join the one above: six countries moved up from the followers to the advanced and fourteen from the marginalized to the followers. Still, this upgrading has been possible for a limited number of economies only, while the majority of countries remained stuck in their own original club. There is therefore the danger that some of them can get into a technological underdevelopment trap – as predicted in Myrdal (1957) cumulative causation – in which the conditions of the past would hamper

---

[23] For a valuable attempt to consider a greater number of elements that differentiate national systems of innovation see, for example, Godinho et al. (2005).

their strategies. This has become a well-established fact for developed economies (see, among others, Patel and Pavitt, 1994; Cantwell, 1991; Vertova, 1998; Andersen, 1998), but which appears to be even more relevant for developing countries.

In particular, as pointed out above, the distance is increasing between the club of marginalized countries and that of followers in relation to the generation of autonomous scientific and technological capabilities. This points to a vicious international division of labour where a few countries generate knowledge for everybody. Even the reduction in literacy rates, the most basic technological capability, is much smaller than for other indicators such as Internet users and telephony. This suggests that policies for this club of countries should be re-thought, integrating national strategies, which often can rely on very limited resources, with the developed policies of the club of advanced countries and with the activities carried out by international organizations.

The analysis carried out here is opening up new research questions. One of them is: how is each club associated with economic performance? Of course, we will assume that the performance of the various members is not homogeneous. But the detailed analysis for a variety of technological sources can possibly identify what is the most appropriate combination of resources to achieve a satisfactory performance. While it is doubtful that the economic policy recipe may be similar for countries that are at the top and at the bottom of the league, it is perhaps possible to single out the good performance achieved by countries belonging to each category. This information could prove vital for the design of appropriate technology policies.

Acknowledgements

Appendix A. The composition of the three technology clubs[24]

Cluster 1: Advanced: Japan, US, Germany, Netherlands, Switzerland, UK, Denmark, Finland, Iceland, Norway, Sweden, Australia, Canada, New Zealand, Israel.

Cluster 2: Followers: Honk Kong (↑), South Korea (↑), Singapore (↑), Malaysia, Philippines, Thailand, Fiji, Austria (↑), Belgium (↑), France (↑), Luxembourg, Cyprus, Greece, Ireland, Italy, Malta, Portugal, Spain, Turkey, Bahrain, Jordan, Kuwait, Lebanon, Saudi Arabia, Syria, United Arab Emirates, Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, Jamaica, Mexico, Panama, Paraguay, Peru, Puerto Rico, Uruguay, Venezuela, South Africa, Trinidad and Tobago, Armenia, Azerbaijan, Belarus, Bulgaria, Croatia, Czech Republic, Georgia, Estonia, Hungary, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Macedonia, Moldova, Poland, Romania, Russian Federation, Slovak Republic, Slovenia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan

Cluster 3: Marginalized: China (↑), Indonesia (↑), Vietnam (↑), Bangladesh, India, Mongolia, Nepal, Papua New Guinea, Pakistan, Sri Lanka, Iran (↑), Oman (↑), Yemen, Albania (↑), El Salvador (↑), Guyana (↑), Honduras (↑), Guatemala, Haiti, Nicaragua, Algeria (↑), Botswana (↑), Mauritius (↑), Tunisia (↑), Zimbabwe (↑), Benin, Cameroon, Central African Republic, Congo Rep., Cote d'Ivoire, Egypt, Gabon, Ghana, Kenya, Lesotho, Madagascar, Malawi, Morocco, Mozambique, Namibia, Nigeria, Senegal, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia.

Appendix B. A test of Durlauf and Johnson's convergence clubs

This appendix presents the results of a multinomial logit test that seeks to analyse the classificatory precision of Durlauf and Johnson's (1995) classification exercise, which is the best-known study in the convergence clubs literature. In their pioneering work, Durlauf and Johnson (D&J) identified four groups of countries by using two discriminatory factors, the GDP per capita and the literacy rate (both taken at their level in 1960) in a large sample of nearly 100 countries (Durlauf and Johnson, 1995, p. 372). The methodology that they used to form the country groupings, the regression tree analysis, estimates a piecewise linear version of the augmented Solow model, thus identifying endogenously the number of clusters to partition the dataset (see Durlauf and Johnson, 1995, p. 381; Breiman et al., 1984).

We have estimated a MNL model for their convergence clubs classification, whose results are presented in Table 5. In this test, the dependent variable is the categorical variable "D&J convergence clubs", which takes value 3 for 'high-output', 2 for 'intermediate output/high-literacy rate', 1 for 'intermediate output/low literacy rate', and 0 for 'low output/low-literacy rate' countries.[25] The explanatory variables are the GDP per capita and the literacy rate, both taken at their level in 1960.[26]

The results confirm the relevance of the D&J convergence clubs classification, and point out its statistical precision in a MNL setting. The estimated coefficients relative to the GDP per capita variable are negative and significant in all the groups, indicating that the higher the

---

[24] The arrows indicate those countries shifting towards the cluster above between 1990 and 2000.

[25] For a list of the countries included in each group, see Durlauf and Johnson (1995, Table 4, p. 374).

[26] The data used for our MNL model estimation have been taken from Durlauf and Johnson (1995)'s 'Data Appendix' (pp. 379–380). D&J's data have been standardized before entering them in the MNL regressions, to be able to compare the results with those relating to our technology clubs classification presented in Tables 4a and 4b, where the explanatory variables were also entered in standardized form.

Table 5
Results of the MNL test for Durlauf and Johnson's 'convergence clubs' (Wald statistics between parentheses)

	Cluster 2: Intermediate-output/high-literacy	Cluster 3: Intermediate-output/low-literacy	Cluster 4: Low-output/low-literacy	Likelihood ratio test
Literacy rate	3.16 (7314.78)[***]	−7.63 (4594.74)[***]	−5.75 (2728.84)[***]	79.54[***]
GDP per capita	−6.02 (9876.00)[***]	−4.98 (3516.48)[***]	−6.62 (6892.79)[***]	44.98[***]
Pseudo R^2	Cox and Snell: 0.839	Nagelkerke: 0.895	R_{LA}: 0.648	
Percentage of correctly predicted cases	Cluster 2: 81.5%	Cluster 3: 88.2%	Cluster 4: 21.4%	Overall: 79.2%

Dependent variable "D&J convergence clubs": $Y = j$, where $j = 3$ for 'high-output', $j = 2$ for 'intermediate output/high-literacy rate', $j = 1$ for 'intermediate output/low literacy rate', and $j = 0$ for 'low output/low-literacy rate' countries.

[***] Significance at the 0.01 level.

GDP per capita the lower the probability that a country is assigned to clusters 2, 3 or 4 instead of group 1 (which is the baseline category).

A similar pattern holds for the literacy rate variable, which is also negatively related to the likelihood that a country is classified into clusters 3 and 4 instead of the more advanced cluster 1. This is not the case for group 2 though, where the estimated coefficient relative to the literacy rate variable turns out to be positive. The reason for the positive sign may be that the restricted group of advanced economies in cluster 1 has already reached very high levels of the literacy rate, while the follower countries in cluster 2 have still some scope for improvement in this respect. So, an increase in this variable is more relevant for countries in cluster 2 than for those in cluster 1, thus increasing the probability that a country is assigned to the former rather than to the latter group.

The lower part of Table 5 reports the overall explanatory power of the MNL model. The model performance is in general terms quite good, although the precision of D&J's classification exercise turns out to be somewhat lower than that of the club classification presented in this paper. In fact, by comparing the lower part of Tables 4a, 4b and 5, we observe that the pseudo R-squared indexes for D&J model are slightly lower than the corresponding indexes for our club model.

More importantly, the percentage of correctly predicted cases is significantly lower for the former than for the latter classification. In particular, the poorest country club in the D&J classification has a rather low classificatory power (around 21%), meaning that the countries included in this group could have very well been placed in other clubs as well. This result is in line with the recent finding of Hansen (2000), who found the existence of very large confidence intervals for the threshold values of D&J piecewise linear regression model, and particularly for their poorest club. Hansen in fact concluded that "the confidence intervals for the threshold parameters are sufficiently large that there is considerable uncertainty regarding their values, hence concerning the proper division of countries into convergence classes as well" (Hansen, 2000, p. 589).

References

Abramovitz, M., 1986. Catching-up, forging ahead and falling behind. Journal of Economic History 46, 385–406.

Adelman, I., Morris, C.T., 1965. A factor analysis of the interrelationship between social and political variables and per capita gross national product. Quarterly Journal of Economics 79, 555–578.

Andersen, B., 1998. The evolution of technological trajectories 1890–1990. Structural Change and Economic Dynamics 9 (1), 5–34.

Archibugi, D., 1992. Patents as indicators of technological innovation. Science and Public Policy 17, 357–368.

Archibugi, D., Michie, J., 1995. The globalisation of technology: a new taxonomy. Cambridge Journal of Economics 19, 121–140.

Archibugi, D., Coco, A., 2004a. A new indicator of technological capabilities for developed and developing countries. World Development 32 (4), 629–654.

Archibugi, D., Coco, A., 2004b. Measuring technological capabilities at the country level: a survey and a menu for choice. Research Policy 34 (2), 175–194.

Barro, R., Lee, J.W., 2001. International data on educational attainment: updates and implications. Oxford Economic Papers 53 (3), 541–563.

Bell, M., Pavitt, K., 1997. Technological accumulation and industrial growth: contrasts between developed and developing countries. In: Archibugi, D., Michie, J. (Eds.), Technology, Globalization and Economic Performance. Cambridge University Press, Cambridge, pp. 83–137.

Breiman, L., Friedman, J.L., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth, Belmont, CA.

Canova, F., 2004. Testing for convergence clubs in income per capita: a predictive density approach. International Economic Review 45 (1), 49–77.

Cantwell, J., 1991. Historical trends in international patterns of technological innovation. In: Foreman-Peck, J. (Ed.), New Perspectives on the Late Victorian Economy: Quantitative Essays on British Economic History. Cambridge University Press, Cambridge, pp. 37–72.

Carlsson, B., 2006. Internationalization of innovation systems: a survey of the literature. Research Policy 35 (1), 56–67.

Castellacci, F., 2007. Evolutionary and new growth theories: are they converging? Journal of Economic Surveys 21 (3), 585–627.

Castellacci, F. Technology clubs, technology gaps and growth trajectories. Structural Change and Economic Dynamics, in press.

Desai, M., Fukuda-Parr, S., Johansson, C., Sagasti, F., 2002. Measuring the technology achievement of nations and the capacity to participate in the network age. Journal of Human Development 3 (1), 95–122.

Desdoigts, A., 1999. Patterns of economic development and the formation of clubs. Journal of Economic Growth 4, 305–330.

Durlauf, S.N., Johnson, P.A., 1995. Multiple regimes and cross-country growth behavior. Journal of Applied Econometrics 10, 365–384.

Edquist, C., 1997. Systems of Innovation: Technologies, Institutions and Organisations. Pinter, London and Washington.

Fagerberg, J., Verspagen, B., 2002. Technology-gaps, innovation-diffusion and transformation: an evolutionary interpretation. Research Policy 31 (8–9), 1291–1304.

Fiaschi, D., Lavezzi, A.M., 2003. Distribution dynamics and nonlinear growth. Journal of Economic Growth 8, 379–401.

Godinho, M.M., Mendonca, S.F., Pereira, T.S., 2005. Towards a Taxonomy of Innovation Systems. Universidade Tecnica de Lisboa, Mimeo.

Grupp, H., Mogee, M.E., 2004. Indicators for national science and technology policy: how robust are composite indicators? Research Policy 33 (9), 1373–1384.

Hansen, B., 2000. Sample splitting and threshold estimation. Econometrica 68 (3), 575–603.

Hobijn, B., Franses, P.H., 2000. Asymptotically perfect and relative convergence of productivity. Journal of Applied Econometrics 15, 59–81.

Howitt, P., Mayer-Foulkes, D., 2005. R&D, implementation and stagnation: a Schumpeterian theory of convergence clubs. Journal of Money, Credit and Banking 37 (1), 147–177.

Islam, N., 1999. International comparison of total factor productivity: a review. Review of Income and Wealth 45 (4), 493–518.

Islam, N., 2003. What have we learnt from the convergence debate? Journal of Economic Surveys 17 (3), 309–362.

ITU, 2001. World Telecommunication Indicators. International Telecommunications Union, Geneva.

Johnson, P.A., Takeyama, L.N., 2001. Initial conditions and economic growth in the US states. European Economic Review 45, 919–927.

Lall, S., Albaladejo, M., 2001. Indicators of the Relative Importance of IPRs in Developing Countries. UNCTAD, Geneva, http://www.ictsd.org/unctad-ictsd/.

Los, B., 2006. A non-parametric method to identify nonlinearities in global productivity catch-up performance. In: Pyka, A., Hanusch, H. (Eds.), Applied Evolutionary Economics and the Knowledge Economy. Edward Elgar, Cheltenham.

Lundvall, B-Å. (Ed.), 1992. National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning. Pinter Publishers, London.

Masanjala, W., Papageorgiou, C., 2004. The Solow model with CES technology: nonlinearities and parameter heterogeneity. Journal of Applied Econometrics 19, 171–201.

Myrdal, G., 1957. Economic Theory and Under-Developed Regions. Duckworth, London.

Nelson, R.R. (Ed.), 1993. National Innovation Systems: A Comparative Analysis. Oxford University Press, New York and Oxford.

NSF, 2000, 2002. Science and Engineering Indicators. National Science Board, National Science Foundation, Washington, DC.

Papageorgiou, C., 2002. Trade as a threshold variable for multiple regimes. Economic Letters 77, 85–91.

Patel, P., Pavitt, K., 1994. Uneven (and divergent) technological accumulation among advanced countries: evidence and a framework of explanation. Industrial and Corporate Change 3, 759–787.

Pavitt, K., 1988. Uses and abuses of patent statistics. In: A. Van Raan, A. (Ed.), Handbook of Quantitative Studies of Science and Technology. Elsevier, Amsterdam.

Peracchi, F., 2001. Econometrics. Wiley, Chichester.

Pianta, M., 1995. Technology and growth in OECD countries, 1970–1990. Cambridge Journal of Economics 19 (1), 175–187.

Prescott, E.C., 1998. Needed: a theory of total factor productivity. International Economic Review 39 (3), 525–551.

Scott Long, J., 1997. Regression Models for Categorical and Limited Dependent Variables. Advanced Quantitative Techniques in the Social Sciences, vol. 7. Sage Publications.

Temple, J., 1999. The new growth evidence. Journal of Economic Literature 37, 112–156.

Temple, J., Johnson, P.A., 1998. Social capability and economic growth. Quarterly Journal of Economics 108, 965–990.

UNDP, 2001. Human Development Report: Making New Technologies Work for Human Development. Oxford University Press, New York.

UNESCO, 2002. World Education Indicators. UNESCO, Paris, http://www.unesco.org.

UNIDO, 2002. Industrial Development Report 2002–2003: Competing through Innovation and Learning. UNIDO, Vienna.

USPTO, 2002. Registered Patent Database. United States Patent and Trademark Office, Washington, DC, http://www.uspto.org.

Vertova, G., 1998. Technological similarity in national styles of innovation in a historical perspective. Technology Analysis & Strategic Management 10 (4), 437–449.

von Tunzelmann, G.N., 1995. Technology and Industrial Progress: The Foundations of Economic Growth. Edward Elgar, Cheltenham.

World Bank, 2003. World Development Indicators. The World Bank, Washington, DC.