International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

Determinants of Long-term Economic Development: An Empirical Cross-country Study Involving Rough Sets Theory and Rule Induction

Obersteiner, M. & Wilk, S.

IIASA Policy Report
September 1999

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Reihe Transformationsökonomie / Transition Economics Series No. 11

Determinants of Long-term Economic Development:

An Empirical Cross-country Study Involving Rough Sets Theory and Rule Induction

Michael Obersteiner, Szymon Wilk

Determinants of Long-term Economic Development:

An Empirical Cross-country Study Involving Rough Sets Theory and Rule Induction

Michael Obersteiner, Szymon Wilk

Reihe Transformationsökonomie / Transition Economics Series No. 11

September 1999

Institut für Höhere Studien
Stumpergasse 56, A-1060 Wien
Fax: +43/1/599 91-163

Michael Obersteiner
Phone: +43/1/599 91-151
E-mail: oberstei@ihs.ac.at

Szymon Wilk
International Institute for Applied Systems Analysis
Schloßplatz 1
A-2356 Laxenburg, AUSTRIA
E-mail: szymon.wilk@cs.put.poznan.pl

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Abstract

Empirical findings on determinants of long-term economic growth are numerous, sometimes inconsistent, highly exciting and still incomplete. The empirical analysis was almost exclusively carried out by standard econometrics. This study compares results gained by cross-country regressions as reported in the literature with those gained by the rough sets theory and rule induction. The main advantages of using rough sets are being able to classify classes and to discretize. Thus, we do not have to deal with distributional, independence, (log-)linearity, and many other assumptions, but can keep the data as they are. The main difference between regression results and rough sets is that most education and human capital indicators can be labeled as robust attributes. In addition, we find that political indicators enter in a non-linear fashion with respect to growth.

Keywords
Economic growth, rough sets, rule induction

JEL Classifications
C49, F43, O40

Comments

Financial support from the Austrian Federal Ministry of Science and Transport (GZ 308.958/I-III/B/8/98) is gratefully acknowledged.

Contents

Introduction

Theoretical work by *inter alia* Romer (1986), Lucas (1988) and Rebelo (1991) have revived the interest of economists on growth theory and its empirics. A simple indicator of this revival is the number of publications on this topic. For 1987 the Econlit-database finds 341 entries on ‘economic growth’ whereas ten years later 905 entries are registered. There are already a number of excellent summaries on the new theory and the empirics of the economic growth (e.g. Barro and Sala-i-Martin (1995), Durlauf and Quah (1998), Aghion and Howitt (1998)).

The first wave of findings associated with new growth theory led to strong and controversial claims with little empirical back-up. This discontent led to improved techniques and new data, which in turn, generated fresh stylized facts on growth with important implications for theory, which also strongly influenced the economists’ policy recommendations (e.g. see Stiglitz 1998).

Krugman (1998) argues that the new growth theory gave rise to a massive industry of cross-country regressions, but with few exceptions, these regressions were neither been closely tied to theory nor did they provide clear evidence in its support. The range of methods and empirical findings related to studying cross-country growth is extensive and continues to grow. However, the more results are generated the more inconsistencies and knowledge (data) gaps are discovered. Durlauf and Quah (1998) feel that the new empirical growth literature still remains in its infancy.

According to Sala-i-Martin (1997) the problem faced by empirical economists is that growth theories are not explicit enough about what variables belong to the “true” regression. That is, even if we know that the “true” model has a certain form, we do not know exactly what variables we should use. A good theorist could make almost any variable affect the level of technology (in a broad sense) and, as a result, the theorist could make almost any variable look like an important theoretical determinant of the rate of economic growth. On the other hand an experienced econometrician could after ‘playing’ with the regressor matrix (e.g. omission of the sub-sahara dummy) tune the results. Thus, new growth theory still lacks effective falsification or verification procedures for theoretical propositions. It becomes obvious that in a world flooded with theoretical propositions on factors influencing long-term economic growth a large variety of methods should be used for testing the validity of these propositions.

Despite the fact that innovation and Schumpeterian entrepreneurship plays a crucial role in the new growth theory, the economic profession has so far relied almost exclusively on the standard econometric toolbox, which reaches its limit in an application where the number of potential regressors exceeds the number of observations. Our innovation is to introduce a new method for investigating the validity and robustness of the theorists’ propositions. In addition, we would like to contribute to a more informed policy discussion by providing a more heterogeneous picture of information embedded in the available growth data. Furthermore, we provide a methodology that helps to identify interesting outliers that might be worth studying. Here we argue that informative policy advice should not only try to ensure obeyance to the robust rules, which are deducted from regression fitting a line through the average, but should in addition provide knowledge gained from frontier, model countries or regions.

In the search for sound empirical findings researchers concentrated on comparing different econometric specifications and implications of growth models. The empirical literature focuses on understanding growth exclusively in terms of factor inputs. It freely uses all kinds of auxiliary explanatory factors in order to distill a set of ‘robust’ determinants. However, using the standard econometric tool kit the researcher puts a lot of *a priori* structure on the system analyzed. Issues of multicollinearity, distributional assumptions of the variables used, poolability, and (log-) linearity are, among others, the main assumptions. These assumptions are not at all innocent with respect to what they are taken for by theorists and in policy discussion. Contrarily, with rough sets we need to make only one assumption, which is that we can form classes. Various class definitions can then again be subject to sensitivity analysis.

This paper provides a description of the proposed method. Since the two methods have never been applied to economic problems before, we first provide a more detailed description of the two methods. Second, we provide a detailed verbal description of our results. Third, we compare our results with those gained from the regression analysis and stick to the structure provided by Sala-i-Martin (1997). Finally, we conclude with a more general discussion on the major new insights gained from our analysis, which we regard as complementary to the existing models used in the empirical literature on economic growth.

Method

Rough Sets Theory

Introduction

The rough sets theory introduced by Pawlak (Pawlak 1982; Pawlak 1991) is a mathematical tool that deals with vagueness and uncertainty. The theory is founded on the assumption that some information is associated with every object of the considered universe (data, knowledge), e.g. the universe considered in this paper consists of countries described by the set of variable values. Objects characterized by the same information are indiscernible (similar) in view of the available information about them. The indiscernibility relation defined in this way constitutes the mathematical basis of the rough sets theory.

A set of indiscernible (similar) objects is called elementary set and forms the basic granule (atom) of knowledge about the considered universe. These granules are used to express concepts about the universe. A concept is defined by the set of representative objects from the universe, e.g. in the considered data set the concept of industrialized, fast-growing country is defined by the set of all such countries. If any concept of the universe can be formed as a union of some elementary sets, it is referred to as crisp (precise). On the contrary, if the concept cannot be presented in such a way, it is referred to as rough (imprecise, vague).

Due to the granularity of knowledge some objects cannot be discerned and appear to be the same (or similar). As a consequence, vague concepts, unlike precise concepts, cannot be characterized in terms of information about their elements. Therefore in the rough sets methodology it is assumed that any vague concept is replaced by a pair of precise concepts – called the lower and upper approximations of the vague concept. The lower approximation consists of all objects which belong to the concept for sure, and the upper approximation contains all objects which possibly belong to the concept. The difference between the upper and the lower approximation constitutes the boundary region of the vague concept.

The basics operations of the rough sets theory are used to discover fundamental patterns in data. Thus, in a certain sense, the rough sets methodology refers to machine learning, knowledge discovery and statistics. However, an interpretation of the obtained results lies outside the theory and can be used in many ways.

Basic concepts

Information system

Information system is the 4-tuple $S=\langle U,Q,V,f\rangle$, where *U* is a finite set of objects (the universe), *A* is the finite set of attributes (features), $V=\bigcup_{q\in A} V_q$ and V_q is a domain of attribute *q* and $f:U\times Q\rightarrow V$ is a total function such that $f(x,q)\in V_q$ for every $q\in Q$, $x\in U$ called *information function*.

Let $S=\langle U,Q,V,f\rangle$ be an information system and let $P\subseteq Q$ and $x, y\in U$. Objects *x* and *y* are *indiscernible* by the set of attributes *P* in *S* iff *f*(x, *q*) = *f*(*x*, *q*). Thus every $P\subseteq Q$ generates a binary relation on *U* which is called *indiscernibility relation* and denoted by *IND*(*P*). The *IND*(*P*) is an equivalence relation for any *P*. Equivalence classes of *IND*(*P*) are called *P-elementary sets* in *S*. The family of all equivalence classes of relation *IND*(*P*) on *U* is denoted by *U* | *IND*(*P*) or, in short, *U* | *P*.

$Des_P(X)$ denotes a description of *P*-elementary set $X\in U \mid P$ in terms of values of attributes from, i.e. $Des_P(X)=\{(q,v): f(x,q)=v, \forall x\in X, \forall q\in P\}$

Approximation of sets

Let $P\subseteq Q$ and $Y\subseteq Q$. The *P-lower approximation* of set (concept) *Y*, denoted as $\underline{P}Y$ and the *P-upper approximation* of *Y*, denoted as $\overline{P}Y$, are defined as:

$$\underline{P}Y=\bigcup\{X\in U\mid P: X\subseteq Y\}$$
$$\overline{P}Y=\bigcup\{X\in U\mid P: X\cap Y\neq\varnothing\}$$

The *P-boundary* (doubtful region) of set *Y* is defined as:

$$Bn_P(Y)=\overline{P}Y-\underline{P}Y$$

Set $\underline{P}Y$ is the set of all elements of *U*, which can be certainly classified as elements of *Y*, employing the set of attributes *P*. Set $\overline{P}Y$ is the set of elements of *U*, which can be possibly classified as elements of *Y*, using the set of attributes *P*. The set $Bn_P(P)$ is the set of elements which cannot be certainly classified to *Y* using the set of attributes *P*.

With every set $Y\subseteq U$, we can associate an *accuracy of approximation* of set *Y* by *P* in *S*, or shortly, accuracy of *Y*, defined as:

$$\boldsymbol{a}_P(Y) = \frac{|\underline{P}Y|}{|\overline{P}Y|}$$

Approximation of classification of U

Let *S* be an information system , $P \subseteq Q$ and $\boldsymbol{Y} = \{Y_1, Y_2, \dots Y_n\}$ be classification of *U*. The origin of the classification is independent on attributes from P, it can be imposed by an export. Subsets Y_i, $i = 1, \dots n$ are categories of classification *Y*. The *P*-lower and *P*-upper approximations of **Y** are defined as sets [lower approximation set] and [upper approximation set], respectively. The coefficient

[quality of approximation of classification]

is called the *quality of approximation of classification* ***Y*** be a set of attributes, or in short, *quality of classification*. It expresses the ratio of all *P*-correctly sorted objects to all objects in the system.

Reduction of attributes

The set of attributes $R \subseteq Q$ *depends* on the set of attributes $P \subseteq Q$ in *S* (denotation $P \rightarrow$ R) if $IND(P) \subseteq IND(R)$. Discovering dependencies between attributes is of primary importance in the rough sets approach to knowledge analysis.

Another important issue is the reduction of attributes. The reduction should lead to the reduced set of attributes that provides the same quality of sorting as the original set of attributes. The minimal subset $R \subseteq P \subseteq Q$ such that $\gamma_P(\boldsymbol{Y}) = \gamma_R(\boldsymbol{Y})$ is called ***Y***-reduct of *P* (or reduct if there is no ambiguity in the understanding of *Y*) and denoted as $RED_g(P)$. The information system may have more than one ***Y***-reduct. Intersection of all ***Y***-reducts is called ***Y***-core of *P*, i.e. [core]. The core is the collection of the most significant attributes in the system.

Decision tables

An information system can be treated as a *decision table* assuming that $Q = C \cup D$, and $C \cap D = \emptyset$, where *C* are called *condition attributes*, and *D* – *decision attributes*.

Decision table $S = < U, C \cup D, V, f >$ is *deterministic* if $C \rightarrow D$; otherwise it is *non-deterministic*. The deterministic decision table uniquely describes the decisions to be made when some conditions are satisfied. In case of a non-deterministic table, decisions are not uniquely determined by the conditions. Instead, a subset of decisions is defined which should be taken into consideration when the conditions are satisfied.

From the decision table a set of decision rules can be derived. Let $U \mid IND(C)$ be a family of all C-elementary sets called *condition classes*, denoted by X_i, $i = 1, \dots, k$. Let $U \mid IND(D)$ be the family of all D-elementary sets called *decision classes*, denoted by Y_j ($j = 1,\dots,n$).

$Des_C(X_i) \Rightarrow Des_D(Y_j)$ is called (C, D)-decision rule. The rules are logical statements "*if…, then…*" relating descriptions of condition and decision classes. The set of decision rules for each decision class Y_j is denoted as $\{r_{ij}\}$. Precisely

$\{r_{ij}\} = \{ Des_C(X_i) \Rightarrow Des_D(Y_j): X_i \cap Y_j \neq \emptyset, i = 1, \dots, k \}$

Rule r_{ij} is *exact* iff $X_i \subseteq Y_j$, and r_{ij} is *approximate* otherwise.

Data preprocessing

Introduction

The classical rough sets theory is based on the indiscernibility relation, which is not suited for real-valued attributes. The use of equality tests in the indiscernibility relation implies that two slightly different values are treated as distinct ones, although from the point of view of the decision maker they are seen as indiscernible ones. Such a property of the indiscernibility relation can lead to an enormous granulation of the data – each object can constitute a single elementary set. This introduces the need for discretization of continuous values.

Recently some extensions of the rough sets theory were introduced that overcome the problem of continuous values without discretization. They introduce similarity relation that replaces indiscernibility relation (see Krawiec *et al.* 1996, Greco *et al*. 1998a). Unfortunately the analysis had already been in progress and these extensions were not taken into consideration.

The indiscernibility relation does not deal with missing values of attributes either. Objects with missing data should be excluded, or the missing values have to be replaced by the known ones. Recently the problem of missing values had been solved without the requirement of modification of original data (see Greco *et al. 1998a,* Greco *et al.* 1998b) but, as in the case of the before-mentioned extensions, it was too late to integrate them into the analysis.

In the presented analysis the classical approach was used: the continuous values were discretized and the missing values were replaced by the known ones.

Discretization of continuous values of attributes

In the presented data analysis process, values of continuous attributes were discretized using a *recursive minimal entropy partitioning* algorithm (Fayyad, Irani 1993). The algorithm represents a *supervised*, *local* and *dynamic* approach to discretization (Dougherty *et al.* 1995). Supervised means that information about classes of objects is used during discretization, *local* means that only one attribute is considered per class, and dynamic –algorithm determines the number of intervals, into which the domain of attribute is divided itself.

For given attribute a the algorithm determines a partition boundary b_{min} that minimizes the class information entropy over all possible boundaries. Partition boundary b is the value of attribute A that divides the domain of A, observed in the set of objects es, into two intervals; this division is reflected by partitioning es into two disjoint subsets es_1 and es_2, considering b and values of A of objects. The class information entropy is defined as:

$$E(A, b, es) = \frac{|es_1|}{|es|} \cdot Entropy(es_1) + \frac{|es_2|}{|es|} \cdot Entropy(es_2),$$

where A is a continuous condition attribute, b is a partition boundary, es is a set of examples, es_1 and es_2 are disjoint subsets of es.

Then the algorithm divides both partitions es_1 and es_2 induced by $b_{\min}$ in the way described above. The recursive partitioning stops when a stopping condition of *Minimal Description Length Principle* is met. The stopping condition is defined as:

$$Gain(A, b, es) < \frac{log_2(|es| - 1)}{|es|} + \frac{\Delta(A, b, es)}{|es|},$$

$$Gain(A, b, es) = Entropy(es) - E(A, b, es),$$

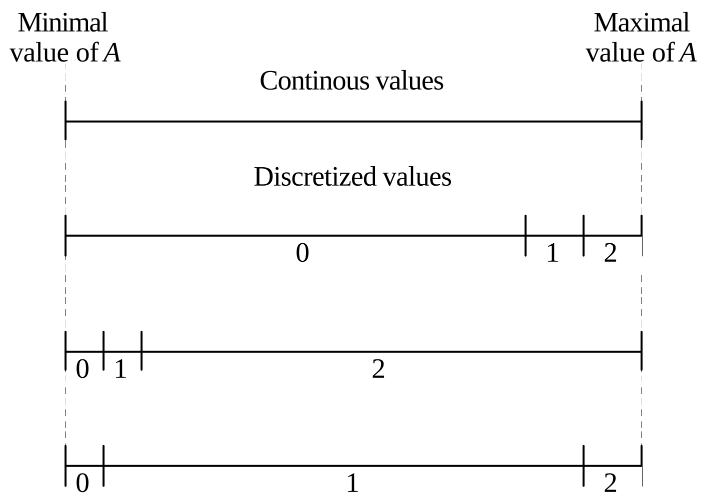
$$\Delta(A, b, es) = log_2(3^k - 2) - \left[k \cdot Entropy(es) - k_1 \cdot Entropy(es_1) - k_2 \cdot Entropy(es_2)\right],$$

where k, k_1, k_2 are the numbers of different classes, to which objects from es, es_1 and es_2 respectively (e.g. if es contains examples from one class, then k is equal to 1) belong.

One should be careful with interpreting the discretized values. For example, if the values of attribute A are discretized into 3 intervals, they are usually coded with 0, 1, and 3, but it can not be taken for granted that 0 corresponds to small values, 1 to medium values, and 2 for large ones. Such situations as these presented on the figure 1 are possible.

Figure 1

Sample discretizations



Processing missing values

Missing values were removed from the data set after the discretization phase. For each condition attribute, for each decision class the most frequent values were identified. If for an object the value of a considered attribute was missing, it was replaced by the most frequent value in the objects class.

Decision rules and their induction

Decision rules

Decision rule *R* is the logical statement of the form **if** *C*, **then** *D*, where *C* is the condition part of a rule, and *D* is the decision part. The condition part is a conjunction of *q* elementary conditions (i.e. $C = c_1 \wedge c_2 \wedge \ldots \wedge c_q$), where elementary condition is a basic test on the values of a condition attribute (e.g. attribute = value). The decision part indicates the decision class *K,* to which objects satisfying the condition part should be assigned.

The set of objects, which satisfy the condition part *C* is denoted as [*C*]. This set can be divided into positive cover $[C]^+_K = [C] \cap K$ and negative cover $[C]^-_K = [C] \cap (U \setminus K)$, where *U* is the set of objects. The decision rules are evaluated mainly by the measure of length *length*(R) = *q*, absolute strength $strength(R) = |[C]^+_K|$, relative strength $strength(R) = |[C]^+_K| / |[K]|$ and the level of discrimination $D(R) = |[C]^+_K| / |[C]|$. The level of discrimination is the probability that an object satisfying the condition part of the rule belongs to the class pointed out in the decision part. If the level of discrimination is equal to 1, then the rule is able to predict exactly the class of the covered object. Otherwise, if the level of discrimination is smaller then 1, then this prediction is only approximate.

Induction of decision rules

As mentioned before it is possible to derive decision rules directly from the decision table using rough sets methodology. However, this can be ineffective as the number of rules may be large (equal to the number of elementary sets) and rules may be too long (they may contain conditions built on all condition attributes). Therefore many algorithms for derivation of decision rules were introduced, e.g. *LEM2* algorithm (Grzymala-Busse, 1992; for a review of algorithms see Stefanowski 1998).

The general procedure of rule induction is presented in figure 2. The algorithm in one run induces rules for a given *concept*. Concept is defined by the set of positive objects (e.g. objects that belong to the concept) and the set of negative objects (e.g. objects not belonging to the concept). In most cases, a concept is identified with a decision class, but other definitions of concept are also possible.

Figure 2

The general rule induction algorithm

```
procedure InduceRules(concept)

rule_set = {}
pos = set of positive examples for concept
neg = set of negative examples for concept
while (pos is not empty)do
     find the best conjunction of conditions ce
     that is satisfied by at least one object from pos and is not
     satisfied by any object from neg
     rule = 'if ce then concept'
     rule_set = rule_set ∪ {rule}
     remove from pos all examples covered by rule
return rule_set
```

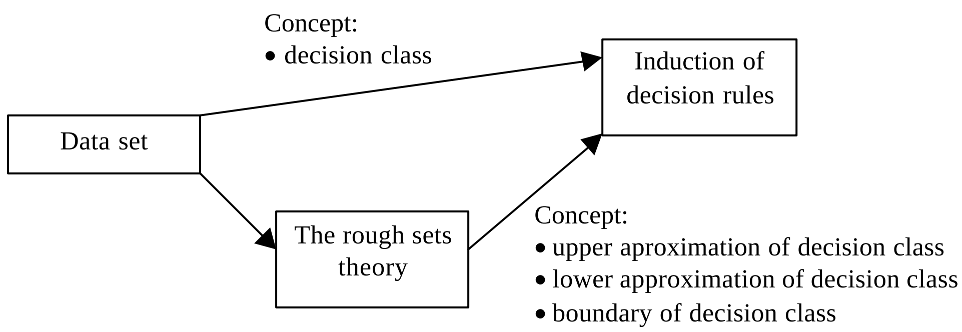
The algorithm showed on figure 2 presents the so-called *separate-and-conquer* approach. In this approach, objects covered by the generated rule are removed from the considered set of positive objects. The algorithm stops if there are no more positive objects to consider. Such a formulation of algorithm ensures that each positive example is covered by exactly one rule.

Another approach, in which objects are not removed from the positive set, is also possible. In such an approach, the stopping condition (line 4) is modified, and usually it contains some heuristics. Unlike in the previous approach, here one object can be covered by many rules. The *Explore* algorithm described below represents such an approach to rule induction.

Here it should be stressed that rule induction is not a part of rough sets theory. It can rather be seen as a tool for preparing data for induction, especially defining concepts, for which rules are generated. For example, rough sets may be used for determining approximations or boundaries of a decision class, which in turn may be used as concepts for rule induction algorithms. The cooperation between rough sets theory and rule induction is presented on figure 3.

Figure 3

The rough sets theory and rule induction



Explore **algorithm**

The *Explore* algorithm (Mienko *et al.* 1996), used for generating rules from the analyzed data set, represents a *discovery oriented* approach to rule induction. Unlike the *classification-oriented* approach, the achievement of the highest possible classification accuracy when classifying new objects is not the main goal, but to extract common patterns and regularities from the analyzed data.

In the algorithm *Explore,* the search for rules is controlled by the parameters called *stopping conditions SC* that reflect the user's requirements. As the main attention is put on the strength of the rules (either absolute or relative), the definition of *SC* is connected with determining the threshold value for the minimal strength of the conjunction that is candidate for the condition part of the rule. If its strength is lower than *SC*, it is discarded, otherwise it can be further evaluated. The *SC* also contains optional bound on maximal conjunction length. If the length is greater then specified in *SC*, then the conjunction is discarded. Additionally, one can define a threshold d expressing the minimum value of the level of discrimination $D(R)$ of the rules to be generated.

The algorithm *Explore* is based on the *breadth-first* strategy, which generates rules of increasing size starting from the shortest ones. The main part of the algorithm is presented in pseudo-code on figure 4.

Figure 4. The *Explore* algorithm

```
procedure Explore(input  LS: list of valid elementary conditions,
                         SC: stopping conditions,
                  output var R: set of rules)
begin
     R ← ∅;
     IsGoodCandidate(LS, R);
          { LS is a list of valid elementary conditions
            c1, c2, … cn – ordered according to the decreasing
            strength }
     Q ← LS;  { Copy current LS to queue Q }
     while Q <> ∅ do
          begin
               select the first conjunction C from Q;
               Q ← Q \ { C };
               generate the set LC of all the conjunctions
                   C ∧ ch+1, C ∧ ch+2, … C ∧ ch+n,
                   where h is the highest index of the condition
                   involved in C
               { generate extensions of C using LS }
               IsGoodCandidate(LC, R);
               Q ← Q ∪ LC;
                   { Place all candidates from LC at the end of Q }
          end
end
procedure IsGoodCandidate(var L: list of conjunctions,
                          var R: set of rules)
{ This procedure prunes list L discarding:
  - conjunctions which are too long according to SC
  - conjunctions whose extension cannot give rise to rules due to SC
  - conjunctions corresponding to rules which are already stored in R
}
begin
     for each C in L do
     begin
          if C satisfies SC then L ← L \ { C }
          else if |[C]K+| / |[C]| ≥ d then
               begin
                    R ← R ∪ { C };
                    L ← L \ { C };
               end
     end
end
```

Data

Data stem from Sala-i-Martin (1997) and relate to variables that were publicly available and appeared before in the growth literature. These variables were usually treated as state variables in the theoretical economic growth literature. Table 1 shows the variables and variable names that were used in the study. This list was not extended with the intention to compare the results of the econometric approach according to Sala-i-Martin (1997) with the results gained from the rough-sets / rule induction approach.

The cross-country panel covers 136 countries. The entire set of countries was subdivided into six disjoint classes. The classes were established on the basis of the IMF classification of industrialized and non-industrialized countries. Industrialized countries were subdivided into slow and fast growing. Non-industrialized were subdivided into slow, medium slow, medium fast, fast growing.

Results

Discretization

The values of continuous variables were discretized using a recursive minimal entropy partitioning algorithm. The discretization results are presented in Table 2. It can be seen that a number of indicators show a high degree of class information entropy, so that the total range of the variable falls into one interval. These variables can be regarded as non-robust variables based on the application of the recursive minimal entropy partitioning algorithm. The list of non-robust variables includes, among others, the average inflation rate, measure of outward orientation and trade openness, public investment share, expenditures of education, wars, political instability and number of revolutions and coups, ethno-linguistic fractionalization and foreign language abilities, real exchange rate distortions, foreign trade growth, share of workers in the population, type of colonialization, religion (except the Confucian, Protestant or Muslim), scale effect (labor force and area), and the fraction of mining.

The symbols **[** and **]** denote the closed (inclusive) left and right boundary; **(** and **)** denote the open (exclusive) left and right boundary. The intervals are numbered from 0 to # Intervals –1. If the continuous value of the attribute belongs to interval no. 1, than it is replaced by 1.

Rough sets

Reducts are computed upon a reduction of all redundant attributes in order to obtain the minimum set of attributes ensuring full approximation of all classes. In other words, the set of attributes contained in the reducts can be interpreted as robust variables explaining our country classes i.e. growth categories. However, there is one qualifier that is the entire set of reducts, and not necessarily the individual attribute, satisfying full approximation. Thus, the result must be interpreted so that only the combination of all attributes leads to growth or stagnation. This makes it difficult to conduct a direct comparison with the regression results, where different variables get different weights expressed by the elasticity. In the rough sets analysis each variable is equally important and necessary to allow full approximation. In other words, our list of attributes is sufficient for serving as a check list to ensure steady economic growth of a country.

We found 264835 reducts. From that huge number of reducts we selected the set of the shortest reducts, namely those containing 11 attributes. There were 277 such reducts. From this reduced set of reducts we filtered those attributes with the largest number of appearances in all reducts. The top 10 of the most frequent attributes were:

- *log(GDP) in 1960(X1)*: 265835;
- Degree of Capitalism(X60): 264835;
- Non-Equipment investment(X43): 174232;
- Public investment share(X24): 163873;
- Domestic credit 60-90(X8): 155029;
- Higher education enrollment*(X22)*: 148507;
- Equipment investment (*X42)*: 139322;
- Years of open economy *(X23)*: 138673;
- Government expending share(*X25)*: 135584;
- Political rights (*X32)*: 130676.

We restricted the reduced set of reducts (277) such that they had to contain at least the 5 most frequent attributes. We found 10 such reducts, which are presented in the Table 6. During the search phase all missing values were treated as unique values i.e. missing values were treated as different from any other known value and there was no distinction made among missing values.

The rough sets analysis was parameterized such that the quality of classification of the data set was equal to 1, i.e. the set was not roughly approximated but with a crisp set. There was no need to calculate a lower and upper approximation of classes in order to induce rules. Thus, the rules were induced directly from the decision classes.

Comparing the results gained from the rough sets analysis with the regression results of Sala-i-Martin (1997) we find that (non-) equipment investment, number of years open economy, political rights, degree of capitalism, public investment share (for model 2 only), and GDP in 1960 (fixed variable in model 1) are consistently labeled as robust variables irrespective of the models used.

Rule induction

The effect of attributes is dependent on the defined class. Different determinants play a different role in each class. Based on our analysis this class-dependency can shortly be described as follows (see results of the rule induction method in Table 2):

Industrialized fast growing countries

The *Explore* algorithm divided, under the parametrization of a maximal rule length of 6 and minimal rule strength of 7, the class ‘industrialized fast growing countries’ into two overlapping sets of countries. The first class is made up of a set of countries mainly including northern European countries such as Finland, France, Island, Belgium, Austria, Norway and Italy. This class of industrialized fast growing countries is described by high equipment investments, a high rate of liquid liabilities to GDP, a medium level of GDP in 1960, low levels of the black market premium and its standard deviation, a long history of open economy, small degree of tariff barriers, high degree of rule of law and democratic freedom, slow population growth, high levels of primary and higher school enrollment, high levels of average years of higher education of total population in 1960, and finally they are all geographically located on higher latitudes. The second class of mainly Mediterranean countries includes Finland, France, Iceland, Italy, Spain, Greece and Portugal. This class however, is characterized by a different, overlapping set of rules. This second class is characterized by high non-equipment investment, high share of domestic credits over the period 1960 to 1990, and a high share of liquid liabilities to GDP. On the institutional side these countries show a low black market premium, an extended period of open economy and a low degree of tariff barriers, and high levels of political rights and civil liberties (inverse scaling of these two variables). The rate of population growth is low, and secondary and tertiary school enrollment is high. The share of Confucians and Muslims is low in this country group. All countries covered are located on higher latitudes.

Japan and Ireland belong to the class of industrialized fast growing countries, but these two countries were not covered by any of the rules. These two countries need to be treated as special cases or outliers, what they essentially are.

Industrialized slow-growing countries

Industrialized slow-growing countries are New Zealand, Switzerland, Australia, Sweden, the United States, Great Britain, Luxembourg, Denmark, the Netherlands, Germany and Canada. All countries belonging to this class are covered by the rule, thus we find 100% coverage. Industrialized slow-growing countries are characterized by a high level of GDP in 1960, high equipment investments, wide coverage of primary education, high life expectancy, high political standards, a capitalist type of production, stable inflation rates and a low black market premium.

Non- industrialized fast growing countries

Non-industrialized fast growing countries show the relatively high coverage of 75%, taking into account all three rules that were generated. Three different rules were generated for the eight, mostly small economies that are represented by this class. Cyprus, Malta, Hong Kong, Taiwan, Singapore and Korea were covered by the rules generated. The first rule covers the above mentioned countries except Korea and describes this class by high levels of equipment investment combined with a low level of GDP in 1960, high levels of secondary and higher education (but low level of average years of primary school education), high level of average years of education of the total population (and its product with log of GDP per capita), low civil liberties and little democratic freedom, high level of domestic credit and its standard deviation, and finally a high black market premium. All countries are geographically located at low latitudes. The second rule covers the above mentioned countries except Hong Kong and it describes the class just by high levels of equipment investment, low democratic freedom and the variables describing characteristics of the population are the same as for the first rule. The third rule covers all above mentioned countries except Singapore and is characterized by the same set of attributes as the first rule, except that the level of domestic credit, the standard deviation of inflation, average years of education of the total population in 1960 and the latitude are not part of the rule. On the other hand the attribute average years of higher education in 1960 appears in the third rule compared to the first rule. Thus, the high growth of Singapore can not be explained by high levels of higher education in 1960.

Botswana and Thailand were not covered by the rules, which indicates that these countries showed a rapid economic growth which must be explained by a different constellation of variables.

Non-industrialized medium fast-growing countries

The class of non-industrialized medium fast growing countries seem to be even more heterogeneous which is illustrated by a coverage of only 27%. There are only five countries covered by the rule, Panama, Ex-Yugoslavia, Egypt, Israel, and Syria. These countries started from a low initial GDP level, they showed high non-equipment investment, a high level of domestic credits, and interestingly they feature a high level of higher education enrollment. Concerning the institutional indicators, these countries can be described as open economies, however with a high degree of tariff barriers, and a low level of the rule of law and civil liberties. For a long list of countries no rule could be found under the chosen parametrisation. This list of countries covers important countries including Brazil, Indonesia, Malaysia, Pakistan, Turkey, Swaziland, Mauritius, Marocco, Barbados, Cap Verde, Tunisia, Seychelles, and Lesotho.

Non-industrialized medium slow-growing countries

Non-industrialized medium slow growing countries are also very heterogeneous, which is reflected in a rather low coverage (45%) of the rule generated. This class is the largest and contains about 36 countries. The countries that are covered by the rule are Surinam, Rwanda, Malawi, Guinea-Bissau, Gambia, South Africa, Nigeria, Togo, Philippines, Bolivia, Chile, Bangladesh, Nepal, Fuji Islands, Paraguay, and Costa Rica. This group is rather diverse with respect to institutional, religious, and population indicators. However, these countries share the following common economic features such as a medium level of equipment investments and, in contrast to the non-industrialized medium high growing class, a low level of non-equipment investment, which might be due to the low levels of higher education, high degree of the level and standard deviation of domestic credit, again in contrast to the non-industrialized medium high growing class a high level of the variability of the inflation rate, high government expenditure share and low levels of the liquid liabilities to GDP ratio. However, these countries were large exporters in 1970. All countries covered by this rule did not show a long history of open economy.

Due to the low coverage we also present results for a weaker parametrization in Table 7.

<u>Non-industrialized slow growing countries</u>

The class of non-industrialized slow-growing countries is mainly made up of Sub-Saharan African countries and covers Chad, Mozambique, Zaire, Uganda, Haiti, Central Africa, Benin, Niger, Burundi, Mauritania, Ghana, Ethiopia, Iraq and Comoros. The coverage of 70% is explained by the countries Madagaskar, Zambia, Nicaragua, Guyana, Mali, and Peru falling out of the rule. This class can be briefly described by the absence of both human and physical capital. This class shows a low level of both equipment and non-equipment investment, which are accompanied by low levels of education at the primary and secondary levels of education.

Comparison

In the following we will compare our results with the regression results of Sala-i-Martin (1997) and Levine and Renelt (1992). Levine and Renelt (1992), applying Leamer's (1983, 1985) extreme bound test, identified only the initial level of income, the investment rate, the secondary schooling enrollment rate, and the rate of population growth as robust empirical attributes of long-term economic growth. Subsequently, a number of researchers concluded that there is nothing to learn from the empirical growth literature. In the following, we will discuss differences of our results with those obtained by Sala-i-Martin (1997).

The fixed variables of Sala-i-Martin (1997)

Log(GDP) in 1960

One of the major controversies of the new growth theory is the proposition that growth patterns of all countries have the tendency to exhibit conditional convergence. This proposition gets strong support from our results for industrialized countries, where at least eighteen out of twenty-three countries show that fast growing industrialized countries started at a medium

GDP level, whereas slow growing countries already started at a high GDP level in 1960. Concerning the classes of non-industrialized countries, we find that all countries that were covered by the rules for non-industrialized fast and medium fast-growing countries started *uni sono* at a low GDP level in 1960. However, only eleven out of eighty-two countries are covered by this attribute. We, thus conclude that the attribute initial GDP is a non-robust variable for classes of non-industrialized countries. Initial GDP is also the most frequent attribute of all reducts. This means that there is a strong indication that the separation into industrialized and non-industrialized countries is justified. This result indicates that pooling across all countries in not justified.

Primary School Enrollment

Primary school enrollment was labeled as non-significant using the extreme bound test used by Levine and Renelt (1992). Sala-i-Martin (1997) concludes that Primary School enrollment is positively but not strongly related to growth. Our analysis shows that primary schooling appears to show up in most of the rules as primary school enrollment and also measured by average years of secondary school education. More interesting, however, is that low levels of average years of primary school education of the total population are observed for non-industrialized fast growing countries. As can be shown for this country class resources for education were concentrated on secondary and for some also on higher education, which explains the spectacular growth pattern. This result is truly interesting and such exceptions from the 'average' rule should receive more attention. There also seems to be a close correlation of low levels of primary education and low levels of civil liberties and democratic freedom.

Life expectancy

In our analysis a high life expectancy – the second fixed variable in Sala-i-Martin (1997) and used there as a proxy for human capital – appears to be a characteristic variable for industrialized countries. In Renelt and Levine (1992) life expectancy was found to be not robust. We conclude that life expectancy is not a explanatory variable of growth between 1960 to 1992, but a result of sustained long-term growth i.e. high living standard.

Other robust variables in Sala-i-Martin (1997)

Investment

Investment appears as equipment and non-equipment investment in every rule induced and also appears as a reduct in the rough sets analysis. This is in line with the results of both Levine and Renelt (1992), and Sala-i-Martin (1997) finds investment a robust variable.

Education

Attributes measuring education and human capital building appear together with investment, as the only most important attributes associated with long-term economic growth. Note that at least one education attribute appears in all rules. High levels of higher and secondary education are attributes of fast growing countries. Our results indicate that education as such is a very robust attribute of economic growth. This stands in contrast to the results if both Sala-i-Martin (1997) and Levine and Renelt (1992), who label education indicators as not strongly related to growth. It appears that higher and secondary education is a necessary condition for high growth independent of the status of being industrialized or not.

Years of open economy

The attribute years of open economy appears for the classes industrialized fast (many years), non-industrialized medium slow and fast growing country classes (few years). This indicates that the attribute 'Years of open economy' can probably be labeled as a robust variable, with the important addition that there appears to be a non-linear relationship of this attribute with growth.

Religious variables

Some religious attributes appear for the description of some classes. They, however, give 'negative' statements that countries of a certain class consist of minorities of one or more confessions. So e.g. industrialized fast growing countries show a small fraction of Muslims. In Sala-i-Martin (1997), however, Muslim has a positive coefficient, and Protestant and Catholic are negative. Therefore it seems that the robustness of these variables might be overestimated in the regression results.

Political variables

Attributes like the rule of law, political rights, civil liberties appear in all rules, except for non-industrialized medium slow- and slow-growing countries. We find that high levels of political indicators are more correlated with the status of being industrialized rather than to growth. Since political attributes do not appear for poor and slow growing countries and political attributes indicate low levels for non-industrialized medium fast and fast growing countries, we unfortunately have to conclude that improvements on political attributes do not lead to improved growth for these countries. As indicated by Barro (1996), there is a non-linear relationship between political variables, which we are able to capture with rule induction. It also appears that political variables are inversely related to education indicators.

Variables that are not strongly correlated with growth in Sala-I-Martin (1997)

Government consumption expenditure

Government consumption expenditure appears as an attribute for fast growing industrialized (low) and medium slow-growing (high) non-industrialized countries. With crisp dicernability of 1 used in the model we conclude that government consumption does not play a major role, but at least in the frontier economies of the fast growing industrialized countries low consumption share seems to be an important qualifier of high growth. Government consumption expenditure also appears in the rough sets approach.

Inflation

Although rough sets would be able to capture a possible non-linear relationship between growth and inflation, inflation, similarly to Sala-i-Martin (1997), measured in average levels does not appear in our decision rules. However, the standard deviation of inflation seems to play a more important role. For industrialized slow-growing and non-industrialized fast growing it appears in low levels and for non-industrialized medium slow growing countries in high levels. Again there is an indication of a non-linear relationship between the variability of inflation and growth.

Openness and other institutional aspects

Openness plays a role in the decision rules generated. The various measures for openness enter in different ways. The measure of outward orientation does not appear in the set of interesting rules, the degree of tariff barriers appears with an indication of low levels for industrialized fast growing countries. The degree of tariff barriers is high in non-industrialized medium fast growing countries. The black market premium and its standard deviation appear with the expected levels for industrialized classes. Interestingly enough, in one rule for non-industrialized fast growing countries the level of the black market premium measure was high. Again, there is some indication of either some non-linear relationship between the black market premium and growth or an indication of interesting outliers.

Population growth

The attribute population growth only appears in decision rules of industrialized countries indicating slow population growth. This is contrary to predictions made by most theoretical growth models and empirical results.

Conclusion

We have provided a new methodology for deriving stylized facts from a data set typical of the new empirics of economic growth. We make a different set of assumptions than those made in the cross-country regression analysis, that is, we can form classes and we are allowed to discretize. Our results are both contradictory and affirmative to those of the regressions performed by Sala-i-Martin (1997). We are able to present a more diverse picture of how different sets of attributes are linked to growth in different classes of countries. Furthermore, we are able to detect non-linearities between an attribute of interest and the growth rate. Rule induction also provides readily available information on interesting exceptions of a generated rule, which can provide valuable information on alternative models to enhance growth (e.g. Japan and Ireland). Our results from both rough sets and rule induction give a strong indication that the cross-section should be partitioned into industrialized and non-industrialized countries. Otherwise, pooling across all countries will result in biased estimates.

From a substantial point of view, our major conclusion is that equipment and non-equipment investment and secondary and higher education are major determinants of high economic growth. The institutional underpinning, as measured by civil liberties and democratic freedom, is not a necessary qualifier of high economic growth, neither is primary schooling.

References

Aghion, Philippe and Peter Howitt (1998): Endogenous Growth Theory. The MIT Press: Cambridge, Mass.

Barro, Robert J. and Sala-i-Martin, Xavier (1995): Economic Growth. McGraw-Hill.

Barro, Robert J. (1996): Determinants of Economic Growth: A cross-country empirical study. *National Bureau of Economic Research Working Paper Series 5698.*

Dougherty, J., Kohavi, R., and Sahami, M. (1995): Supervised and Unsupervised Discretizations of Continuous Features. In: *Proceedings of the 12th International Conference on Machine Learning*, Morgan Kaufmann Publishers, New York, 194-202.

Durlauf, Steven N. and Danny Quah (1998): The empirics of economic growth. NBER - Working Paper Series 6422, February.

Fayyad U.M. and Irani, K.B. (1993): Multi-interval discretization of continuous-valued attributes for classification learning. In: *Proceedings of the 13th International Conference on Artificial Intelligence*, Morgan Kaufmann, pp. 1022–1027.

Greco S., Matarazzo, B., and Slowinski, R. (1998a): On joint use of indiscernibility, similarity and dominance in rough approximation of decision classes. In: *Proc. 5th International Conference of the Decision Science Institute*. Athens (1998).

Greco S., Matarazzo, B., and Slowinski, R. (1998b): Rough set analysis of information tables with missing values. In: *Proc. 5th Internaional Conference of the Decision Science Institute.* Athens (1999).

Grzymala-Busse, J.W. (1992): LERS – A system for learning from examples based on rough sets. In: Slowinski, R. (ed.), *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Set Theory*. Kluwer Academic Publishers, Dordrecht, pp. 3–18.

Krawiec, K., Slowinski, R., and Vanderpooten, D. (1996): Construction of rough classifiers based on application of a similarity relation. In: S.Tsumoto et al., eds. *Proc. of the Fourth Internat. Workshop on Rough Sets, Fuzzy Sets and Machine Discovery – RSFD'96.* Tokyo, 23–30.

Krugman, Paul (1998): Space: The last frontier. *Journal of Economic Perspectives*, Vol. 12, Nr. 2; Spring 1998, pp. 161–174.

Levine, R. and Renelt, D. (1992): “A sensitivity analysis of Cross-Country Growth regressions,” *American Economic Review*, 82.

Lucas Jr., Robert E. (1988): On the mechanics of economic development, *Journal of Monetary Economics*, 22(3):3–42, June.

Mienko, R., Stefanowski, J., Tuomi, K., and Vanderpooten, D. (1996): Discovery-oriented induction of decision rules. *Cahier du LAMSADE* **141**.

Pawlak, Z., (1982): Rough sets. *International Journal of Information and Computer Sciences* 11: 341–356.

Pawlak, Z. (1991): Rough Sets: Theoretical Aspects of Reasoning about Data’, Kluwer Academic, Dordrecht.

Rebelo, Sergio (1991): Long-Run Policy Analysis and Long-Run Growth, *Journal of Political Economy*, 99(3):500–521, June.

Romer, Paul M. (1986): Increasing Returns and Long-run Growth, *Journal of Political Economy*, 94(5):1002-1037, October.

Sala-i-Martin, X.X. (1997): I just ran four million regressions. *National Bureau of Economic Research Working Paper Series 6252.*

Stefanowski, J. (1998): On rough set based approaches to induction of decision rules. In *Rough Sets in Knowledge Discovery Vol 1*, Skowron A, Polkowski, L. (eds); Physica Verlag: Heidelberg; 500–529.

Stiglitz, Joseph (1998): More Instruments and Broader Goals: Moving Toward the Post-Washington Consensus. The 1998 WIDER Annual Lecture, Helsinki, Finland, January 7, 1998.

Table 1 List of attributes used for the analysis

#	Var Name	.
X1	GDPSH60	log(GDP per capita 1960) ref. Summers-Heston, from Barro and Lee 1993
X2	LIFEE060	Live expectency, from Barro and Lee 1993
X3	P60	Primary school enrollment, from Barro and Lee 1993
X4	safrica	Sub-Sahara African Dummy
X5	laam	Latin American dummy
X6	bmp1	Black markt premium (1960-1989), Levine and Renelt 1992
X7	BMS6087	Standard deviation of Black markt premium (1960-1989), Levine and Renelt 1992
X8	GDC6089	Growth of domestic credit 1960-1990, Levine and Renelt 1992
X9	STDC6089	Standard deviation of of domestic credit 1960-1990, Levine and Renelt 1992
X10	PI6089	Average Inflation rate 1960-1989, Levine and Renelt 1992
X11	STPI6089	Standard deviation of Inflation rate 1960-1989, Levine and Renelt 1992
X12	SCOUT	Measure of outward orientation, Levine and Renelt 1992
X13	area	Area (scale effect), Barro and Lee 1993
X14	freeop	Measure of free trade openness, Barro and Lee 1993
X15	freetar	Degree of tariff barriers, Barro and Lee 1993
X16	dpop6090	Growth rate of population, Average rate between 1960-1990, Barro and Lee 1993
X17	pyr60	Average years of primary school education of total population in 1960, Barro and Lee 1993
X18	syr60	Average years of secondary school education of total population in 1960, Barro and Lee 1993
X19	hyr60	Average years of higher education of total population in 1960, Barro and Lee 1993
X20	human60	Average years of education of total population in 1960, Barro and Lee 1993
X21	s60	Secondary school enrollment, from Barro and Lee 1993
X22	h60	Higher education enrollment, from Barro and Lee 1993
X23	YrsOpen	Number of years Open Economy, Sachs and Warner 1996
X24	ggcfd3	Public investment share as fraction of GDP, Barro and Lee 1993
X25	gvxdxe52	Government consumption expenditures (net of defence and education expenditures), Barro and Lee 1993
X26	geerec1	Nominal expenditures for education as a fraction of GDP, Barro and Lee 1993
X27	gde1	Nominal expenditures for defence as a fraction of GDP, Barro and Lee 1993
X28	assassp2	Number of political assassinations, Barro and Lee 1993
X29	revcoup	Number of revolutions and coups, Barro and Lee 1993
X30	pinstab2	Political Instability, Knack and Keefer 1995
X31	wardum	Dummy for countries that have been involved in wars any time between 1960 and1990, Barro and Lee 1993
X32	prightsb	Political rights, Barro 1996
X33	civlibb	Index of Civil liberties, Knack and Keefer 1995
X34	ABSLATIT	Absolute latitude, Barro 1996
X35	AGE	Average age of population, Barro and Lee 1993
X36	FRAC	Ethnolinguistic Fractionalization. Probability two radom people in a country do not speak the same language, Easterly and Levine 1996
X37	DEMOC65	Qualitative index of democratic freedom, Knack and Keefer 1995
X38	PRIEXP70	Fraction of primary exports in total exports in 1970, Sachs and Warner 1996b
X39	RULELAW	Rule of Law, Barro 1996
X40	URB60	Urbanization, fraction of people living in cities, Barro and Lee 1993
X41	RERD	Real exchange rate distortion, Barro and Lee 1993
X42	EQINV	Equipment investment Delong and Summers 1991
X43	NONEQINV	Non-Equipment investment Delong and Summers 1991
X44	humanyl	Product of average years of scholling and log of GDP per capita in 1960, Barro and Lee 1993
X45	tot1	Growth of terms of trade between 1960 and 1990, Levine and Renelt 1992
X46	work60l	Ratio of Workers to population, Barro 1996
X47	lly1	Liquid liabilities to GDP (measure of financial development), King and Levine 1993
X48	BRIT	Dummy variable for former British colony, Barro 1996
X49	FRENCH	Dummy variable for former French colony, Barro 1996
X50	SPAIN	Dummy variable for former Spanish colony, Barro 1996
X51	BUDDHA	Fraction of population that follows Buddist Religion, Barro 1996
X52	CATH	Fraction of population that follows Catholic Religion, Barro 1996
X53	CONFUC	Fraction of population that follows Confucian Religion, Barro 1996
X54	HINDU	Fraction of population that follows Hindu Religion, Barro 1996
X55	JEW	Fraction of population that follows Jewish Religion, Barro 1996
X56	MUSLIM	Fraction of population that follows Muslim Religion, Barro 1996
X57	PROT	Fraction of population that follows

		Protestant Religion, Barro 1996
X58	lforce60	Size of labor force (scale effect), Barro and Lee 1996
X59	Mining	Fraction of GDP in Mining, Hall and Jones 1996
X60	EcOrg	Index of degree in which economies favor capitalist forms of production, Hall and Jones 1996
X61	OthFrac	Fraction of population speaking foreign languages
X62	EngFrac	Fraction of population speaking English as foreign language

Table 2 Discretization results using the recursive minimal entropy partitioning algorithm[1]

Attribute	# Intervals	Intervals
X1	3	[5.517500,8.350901)
		[8.350901,8.706499)
		[8.706499,9.187500]
X2	2	[32.299999,67.500000)
		[67.500000,73.400002]
X3	2	[0.050000,0.905000)
		[0.905000,1.000000]
X4	2	[0.000000,0.500000)
		[0.500000,1.000000]
X5	1	[0.000000,1.000000]
X6	2	[0.000000,0.010500)
		[0.010500,1.916000]
X7	2	[0.001000,0.998050)
		[0.998050,588.626709]
X8	2	[-15.424000,15.029000)
		[15.029000,134.729996]
X9	2	[2.730000,8.656500)
		[8.656500,589.801025]
X10	1	[2.657000,473.779999]
X11	2	[1.760000,6.062000)
		[6.062000,2130.699951]
X12	1	[0.000000,1.000000]
X13	1	[0.500000,9976.000000]
X14	1	[0.078500,0.416200]
X15	2	[0.000000,0.017950)
		[0.017950,0.109900]
X16	2	[-0.000400,0.013700)
		[0.013700,0.038500]
X17	2	[0.053000,4.654500)
		[4.654500,7.704000]
X18	2	[0.008000,0.230000)
		[0.230000,2.872000]
X19	2	[0.000000,0.039000)
		[0.039000,0.530000]
X20	3	[0.072000,5.277000)
		[5.277000,7.513500)
		[7.513500,9.612000]
X21	2	[0.005000,0.195000)
		[0.195000,0.860000]
X22	2	[0.000000,0.033000)
		[0.033000,0.321000]
X23	3	[0.000000,0.578000)
		[0.578000,0.978000)
		[0.978000,1.000000]
X24	1	[0.003700,0.237600]
X25	2	[0.005700,0.087300)
		[0.087300,0.383900]
X26	1	[0.004000,0.053000]
X27	1	[0.002000,0.166000]
X28	1	[0.000000,0.253000]
X29	1	[0.000000,1.190000]
X30	1	[0.000000,0.501700]
X31	1	[0.000000,1.000000]
X32	3	[1.000000,2.416650)
		[2.416650,6.444450)
		[6.444450,7.000000]
X33	2	[1.000000,1.972200)
		[1.972200,6.888900]
X34	2	[0.228000,36.854000)
		[36.854000,63.891998]
X35	1	[0.000000,90.000000]
X36	1	[0.000000,0.930000]
X37	2	[0.072000,0.878000)
		[0.878000,1.000000]
X38	2	[0.041000,0.468000)
		[0.468000,1.000000]
X39	2	[0.000000,0.750000)
		[0.750000,1.000000]
X40	2	[0.015000,0.490500)
		[0.490500,1.000000]
X41	1	[51.000000,277.000000]
X42	3	[0.000200,0.009450)
		[0.009450,0.050850)
		[0.050850,0.148200]
X43	2	[0.029900,0.130150)
		[0.130150,0.280300]
X44	3	[0.461100,45.133148)
		[45.133148,65.540848)
		[65.540848,86.306503]
X45	1	[-0.081200,0.086000]
X46	1	[-1.350900,-0.047000]
X47	2	[0.028800,0.319950)
		[0.319950,1.598600]
X48	1	[0.000000,1.000000]
X49	1	[0.000000,1.000000]
X50	1	[0.000000,1.000000]
X51	1	[0.000000,0.950000]
X52	1	[0.000000,1.000000]
X53	2	[0.000000,0.080000)
		[0.080000,0.600000]
X54	1	[0.000000,0.900000]
X55	1	[0.000000,0.820000]
X56	2	[0.000000,0.045000)
		[0.045000,1.000000]
X57	2	[0.000000,0.435000)
		[0.435000,0.980000]
X58	1	[67.659203,195781.453125]
X59	1	[0.000000,0.533000]
X60	2	[0.000000,3.500000)
		[3.500000,5.000000]
X61	1	[0.000000,1.004000]
X62	1	[0.000000,1.000000]

[1] The names of the attributes refer to Table 1.

Table 3 Decision table

	I_FAST	I_FAST	I_SLOW	NI_FAST	NI_FAST	NI_FAST	NI_MED_HI	NI_MED_LO	NI_SLOW
Economic									
Equipment investment	2		2	2	2	2		1	0
Non-Equipment investment		1					1	0	0
Domestic credit 60-90		1		1			1	1	
S.D. Domestic credit 60-90				1		1		1	
S.D. Inflation 60-90			0	0				1	
Primary exports in 1970								1	
Liquid Liabilities to GDP	1	1						0	
Government expending share		0						1	
Log(GDP) in 1960	1		2	0		0	0		
Institutional									
Black market premium	0	0	0	1					
S.D.Black market premium	0		0						
Degree of Capitalism			1						
Years of open economy	1	1					0	0	
Degree of tariff barriers	0	0					1		
Rule of law	1		1				0		
Political rights		0							
Civil liberties	-		0	1		1	1		
Democratic freedom	1		1	0	0	0			
Population									
Growth rate of population	0	0							
‘Live expectancy	1		1						
Average years of education of total population in 1960				0	0				
Product of average years of schooling and log of GDP per capita in 1960				0	0	0			
Primary school enrollment	1								0
Average years of primary school education of total population in 1960,			1	0	0	0			
Secondary school enrollment, from		1		1	1	1			
Average years of secondary school education of total population in 1960,									0
Higher education enrollment,	1	1					1	0	
Average years of higher education of total population in 1960,	1					1			
Religion									

Fraction of Confucious	-		0		0
Fraction of Protestants					0
Fraction of Muslim		0	0		
Absolute Latitude		1	1	0	

Table 4 List of countries that were covered and not covered by the generated rules

	covered by the rule	not covered by the rule
I_FAST	FIN, FRA, ISL, BEL, AUT, NOR, ITA	ESP, GRC, PRT, JPN, IRL
I_FAST	FIN, FRA, ISL, ITA, ESP, GRC, PRT	AUT, NOR, ITA, JPN, IRL
I_SLOW	NZL, CHE, AUS, SWE, USA, GBR, LUX, DNK, NLD, DEU, CAN	
NI_FAST	CYP, MLT, HKG, OAN, SGP	THA, BWA, KOR
NI_FAST	CYP, MLT, OAN, SGP, KOR	THA, BWA, HKG
NI_FAST	CYP, MLT, HKG, OAN, SGP	THA, BWA,KOR
NI_FAST	CYP, MLT, HKG, OAN, KOR	THA, BWA, SGP
NI_MED_HI	PAN, YUG, EGY, ISR, SYR	SWZ, BRA, MUS, PAK, TUR, MAR, BRB, CPV, TUN, SYC, LSO, IDN, MYS
NI_MED_LO	SUR, RWA, MWI, GNB, GMB, ZAF, NGA, TGO, PHL, BOL, CHL, BGD, NPL, FJI, PRY, CRI	SEN, VEN, CIV, HVO, ZWE, URY, ARG, GTM, HND, IRN, JAM, TTO, DZA, IND, JOR, MEX, COL, ECU, COG, GAB
NI_SLOW	IRQ, TCD, MOZ, ZAR, UGA, HTI, CAF, BEN, NER, BDI, MRT, GHA, ETH, COM	MDG, ZMB, NIC, GUY, MLI, PER

Table 5 Parameters for the rule induction algorithm

	Maximal rule length	Minimal rule strength
I_FAST	6	7
I_SLOW	6	11
NI_FAST	6	5
NI_MED_HI	5	5
NI_MED_LO	6	16
NI_SLOW	6	14

Table 6 Reducts containing the 5 most frequent attributes[2]

1	X1	X8	X11	X18	X21	X24	X25	X42	X43	X56	X60
2	X1	X3	X8	X18	X24	X25	X32	X43	X47	X56	X60
3	X1	X3	X4	X8	X11	X24	X32	X43	X47	X56	X60
4	X1	X3	X8	X11	X17	X24	X32	X43	X47	X56	X60
5	X1	X3	X8	X11	X18	X24	X32	X43	X47	X56	X60
6	X1	X3	X8	X11	X20	X24	X32	X43	X47	X56	X60
7	X1	X3	X8	X11	X24	X32	X43	X44	X47	X56	X60
8	X1	X3	X6	X8	X17	X23	X24	X32	X43	X47	X60
9	X1	X3	X6	X8	X20	X23	X24	X32	X43	X47	X60
10	X1	X3	X6	X8	X23	X24	X32	X43	X44	X47	X60

[2] The names of the attributes refer to Table 1 and the five most frequent attributes are shaded.

Table 7 Overview of weaker decision rules for non-industrialized medium slow growing countries (Minimal rule strength = 8 countries, Maximal rule lenth = 4 conditions)

EQINV	NONEQINV	GDC6089	STDC6089	STPI6089	PRIEXP70	lly1	gvxdxe52	GDPSH60	bmp1	BMS6087	EcOrg	YrsOpen	freetar	RULELAW	prightsb	civlibb	DEMOC65
1	0																
							1			1							
1					1	0						0	1	0	0		
											1	0					
							1										
						0	1										
				1			1		1								
1					1		1		1				1				
				1	1		1		1	1							
							1		1								
							1		1			0					
							1		1					0			
		1		1			1										
1	0		1								1						
1				1			1										
1	0		1								1						
											1		1		1		
	0						1										
							1										
						0	1										
1	0						1										
1							1										
1							1										

pop6090	LIFEE060	human60	humanyl	p60	pyr60	s60	syr60	h60	hyr60	CONFUC	PROT	MUSLIM	abslatit	safrica	Country Class
				1											VEN, SUR, URY, ARG, PHL, CHL, PRY, CRI, GAB
												0			RWA, SLV, PNG, BOL, CHL, NPL, DOM, PRY, CRI, ECU, COG
1	0	0	0		0									0	VEN, JAM, PNG, TTO, IND, DOM, CRI, COL
	0	0	0		0		1								CIV, ARG, JAM, NGA, CMR, CHL, FJI, GAB
				0			1	0							ZAF, NGA, CMR, BGD, TZA, FJI, ECU, COG
				0			1								NGA, CMR, BOL, BGD, TZA, FJI, ECU, COG
												0			SLV, BOL, CHL, NPL, DOM, PRY, CRI, ECU, COG
										0		0			SLV, PNG, BOL, CHL, NPL, DOM, PRY, CRI, ECU, COG
1									1				0		PHL, BOL, CHL, TZA, FJI, PRY, CRI, ECU, COG
						0						0			SLV, PNG, BOL, NPL, DOM, PRY, ECU, COG
												0			] {SLV, PNG, CHL, NPL, DOM, PRY, CRI, COG
												0			SLV, PNG, BOL, NPL, DOM, PRY, CRI, ECU, COG
												0			RWA, BOL, CHL, NPL, DOM, PRY, CRI, ECU
															SEN, CIV, ARG, MWI, GMB, NGA, CHL, NPL, FJI, GAB
												0			RWA, SLV, BOL, CHL, NPL, DOM, PRY, CRI, ECU, COG
															CIV, ARG, MWI, GMB, NGA, CHL, NPL, FJI, GAB
							1								CIV, ARG, NGA, CMR, CHL, FJI, JOR, GAB
1							1								ZAF, NGA, CMR, PHL, BOL, CHL, BGD, FJI, PRY, CRI
						0	1	0							ZAF, NGA, CMR, BGD, TZA, FJI, PRY, ECU, COG
						0	1								NGA, CMR, BOL, BGD, TZA, FJI, PRY, ECU, COG
							1								SUR, ZAF, NGA, PHL, BOL, CHL, BGD, FJI, PRY, CRI
						0						0			RWA, SLV, PNG, BOL, NPL, DOM, PRY, ECU, COG
								0				0			RWA, SLV, PNG, NPL, DOM, PRY, ECU, COG

Institut für Höhere Studien
Institute for Advanced Studies
Stumpergasse 56
A-1060 Vienna
Austria

Phone: +43-1-599 91-149
Fax: +43-1-599 91-163
e-mail: woergoet@ihs.ac.at