R U T C O R
RESEARCH
R E P O R T

COUNTRY RISK RATINGS: STATISTICAL AND COMBINATORIAL NON-RECURSIVE MODELS

P.L. HAMMER[a] A. Kogan[b] M.A. Lejeune[c]

RRR 8-2004, MARCH 2004

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

[a] hammer@rutcor.rutgers.edu , Tel.: +1 7324454812 ; Fax: +1 7324455472
[b] kogan@rutcor.rutgers.edu
[c] mlejeune@andromeda.rutgers.edu
[a,b,c] RUTCOR - Rutgers University Center for Operations Research, Piscataway, NJ, USA
[a,b] Rutgers Business School, Rutgers University, Newark-New Brunswick, NJ, USA

RUTCOR RESEARCH REPORT

RRR 8-2004, MARCH 2004

COUNTRY RISK RATINGS: STATISTICAL AND COMBINATORIAL NON-RECURSIVE MODELS

P.L. Hammer A. Kogan M.A. Lejeune

Abstract. The central objective of this paper is to develop transparent, consistent, self-contained, and stable country risk rating systems, closely approximating the country risk ratings provided by a major rating agency (Standard & Poor). We propose two models that achieve the stated objectives, the first one utilizing the classical econometric technique of multiple linear regression, and the second one using the combinatorial-logical technique of Logical Analysis of Data. The proposed models use economic-financial and political variables, and are non-recursive (i.e., they do not rely on the previous years' ratings). The accuracy of the proposed models' predictions, measured by their correlation coefficients with Standard and Poor's ratings, and confirmed by k-folding cross-validation, exceeds 95%. The stability of the constructed non-recursive models is shown in three ways: by the correlation of the predictions with those of other agencies (Moody's and The Institutional Investor), by predicting 1999 ratings using the non-recursive models derived from the 1998 dataset applied to the 1999 data, and by successfully predicting the ratings of several previously non-rated countries. The confidence in the results and in the validity of both models is strongly reinforced by the fact that the traditional linear regression model and the qualitatively different combinatorial-logical model produce almost identical results.

Acknowledgements: The authors express the appreciation to Dr. Sorin Alexe for his invaluable help in the execution of computational experiments with LAD.

1 Country Risk, Country Risk Ratings and Objectives of the Paper

1.1 Country risk, country risk ratings and their importance

The globalization of the world economies, and in particular the internationalization of financial markets in the last decades, have dramatically expanded and diversified investment possibilities, leading to numerous new opportunities, accompanied by new risks. Consequently, there has been growing interest in obtaining reliable estimates of the risk of investing in different countries. These concerns have led to the development of the concept of country risk, and even to the regular publication of country risk ratings by various agencies. The importance of ratings has been magnified by the recommendations addressed in the Basel Capital Accord (2001), that pinpoints the role of agencies' ratings for the assessment of credit risk.

Different definitions have been proposed for country risk, i.e. for the risk that a country defaults on its obligations. The existing literature on the topic recognizes both financial/economic and political components of country risk. According to the degree to which some of these components are emphasized, country risk is viewed either from the financial/economic perspective only, or from the combined financial/economic and political perspectives.

There are two basic approaches to the interpretation of the reasons for defaulting. The debt-service capacity approach focuses on the deterioration of solvency of a country, which prevents it from fulfilling its commitments. For instance, Bourke and Shanmugam (1990) define country risk as "the risk that a country will be unable to service its external debt due to an inability to generate sufficient foreign exchange". Within this framework, country risk is viewed as a function of various financial and economic country parameters. The cost-benefit approach views a default on commitments or a rescheduling of debt as a deliberate choice of the country, which may prefer this alternative over repayment, in spite of its possible long-term negative effects (e.g. the country's exclusion from certain capital markets (Reinhart,2002), reputation damage). Since the deliberate decision to default results from a political process, political country parameters are included in this type of country risk modeling, along with the financial and economic ones. This approach is strongly recommended by Brewer and Rivoli (1990, 1997) as well as Citron and Neckelburg (1987), who emphasize the impact of the political stability indicator on country risk ratings.

In response to the increased demand for the evaluation of creditworthiness, several agencies such as Moody's, Standard & Poor, Fitch, The Institutional Investor, Euromoney, Dun & Bradstreet, etc. have developed expertise in estimating country risk. These estimates are presented in the form of ratings, or scores, and are generally viewed as indicative of possible future default. Haque et al. (1996) define country credit risk ratings compiled by commercial sources as an attempt "to estimate country-specific risks, particularly the probability that a country will default on its debt-servicing obligations". Sovereign ratings can be viewed as the probability that a borrowing country will fail to pay back.

Country (or sovereign) risk ratings impact countries in a number of ways. The primary significance of ratings is due to their influence on the interest rates at which countries can obtain credit on the international financial markets: the higher the ratings (i.e., the lower the risk of default) the lower the interest rate. Following its sovereign rating downgrade , Japan's borrowing became more expensive as interest rates have increased, reflecting the higher chance of default , which deteriorates even more the situation of the heavily indebted Japanese government and economy.

Second, sovereign ratings also influence credit ratings of national banks and companies, and affect their attractiveness to foreign investors. Ferri et al. (2001) call sovereign ratings the "pivot of all other country's ratings". Similarly, Erb et al. (1995a) underline that raters have historically shown a reluctance to give a company a higher credit rating than that of the sovereign where the company operates. For example, after Moody's downgraded Japan in November 1998 (from Aaa to Aa1), all other Aaa Japan issuers have been downgraded (Jüttner and McCarthy, 2000). This led sovereign ratings to be named "sovereign credit risk ceilings".

Third, institutional investors are sometimes contractually restricted on the degree of risk they can assume, implying in particular that they cannot invest in debt rated below a prescribed level. Ferri et al. (2001) refine this analysis, pointing out the contrast between the ratings of banks operating in high- and low-income countries, and show that ratings of banks operating in low-income countries are significantly affected by variations in sovereign ratings, while the ratings of banks operating in high-income countries do not seem to depend significantly on country ratings. Similarly, Kaminsky and Schmukler (2002) as well as Larrain et al. (1997) note that sovereign ratings are crucial for developing economies, which have a very high sensitivity to rating announcements.

1.2 Critiques of present rating systems

The purpose of ratings is that of compressing a variety of information about a country into a single parameter which can be easily understood, and therefore conveniently used in a decision making process involving comparisons between different countries. Consequently, ratings provide aggregations of diverse indicators into a single metric and can be viewed as a kind of "commensuration" (Kunczik, 2000). The interpretation of ratings is complicated by the heterogeneity of indicators (political stability, inflation, etc.) which may have been used in deriving them.

Comprehensibility: The country risk ratings published by different agencies appear as outputs of "black boxes", the real content and meaning of which are unexplained and hard to understand, since rating agencies specify neither the factors which are taken into consideration in determining their ratings, nor the "rules of compression" of multiple factors into a single rating. This raised the discontent of Japan's Prime Minister, Junichiro Koizumi, who was "railed at being rated in the same neighborhood as African countries to which Japan is providing assistance" . Officials of Japan's Ministry of Finance added that big rating agencies are "making unfair qualitative judgments" , while Moody's denied and claimed that the motives for the downgrade lie in the "increased debt load" of Japan. In view of such controversy, uncovering both the factors which are taken into account by these black boxes, and the mechanisms of deriving ratings, are essential for ascertaining the consistency of a country rating system.

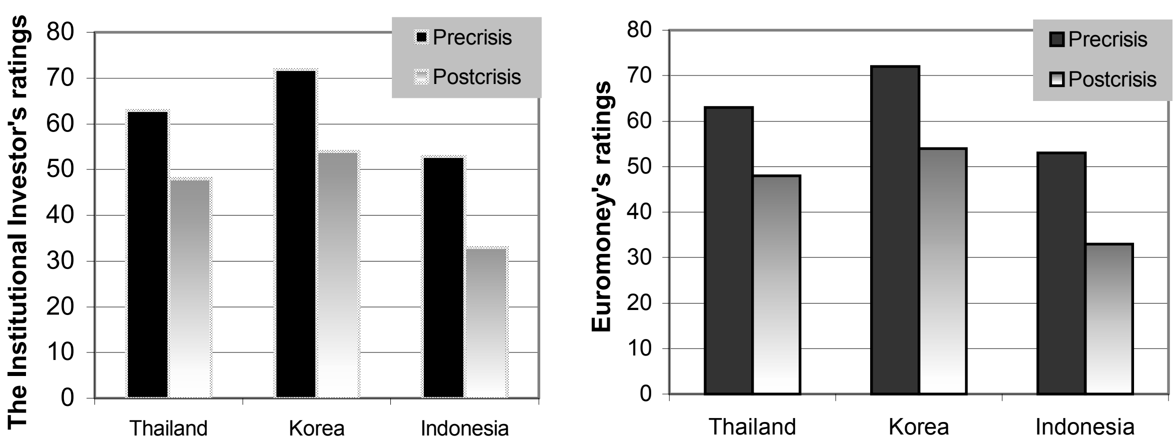
Unknown factors: It is generally assumed that ratings are obtained by aggregating economic/financial and/or political variables. Clearly, the main objective of any country risk rating system is to represent the creditworthiness of countries, i.e., their capacity to pay off loans.

It is not clear however which ones of the many possible factors do actually influence the payback capacity of a country. This question is subject to different analyses. Haque et al. (1998) claim that it is sufficient to restrict the scope of analysis to economic/financial factors only, while others (Brewer and Rivoli,1990) claim that both economic/financial and political factors impact country risk ratings.

Rating failures: Some recent failures have challenged the trustworthiness of country risk ratings (Reinhart, 2002, Levich et al., 2002) . Criticisms directed towards ratings institutions have been especially intense after the Tequila and the Asian crises. Indeed, the tequila crisis in Mexico (1994-95) had not been preceded by a rating downgrade, implying that either the crisis was not predicted, or that its significance was overlooked. Similar observations apply to the Asian crisis (1997-99): Ftich admitted that "it and its larger rivals Standard's & Poor and Moody's Investors Services of the US had largely failed to predict the recent turmoil in Asia" . On the other hand, rating agencies have been more insightful in anticipating other crises, e.g. in Russia (1998), Brazil (1998) and Argentina (2001).

- **Regional bias:** Diverse explanations have been provided for the failure of rating agencies to signal crisis emergencies in various countries. There are claims that certain rating agencies favor certain regions. For instance, Haque et al. (1997) note that Euromoney usually gives higher ratings to Asian and European countries than to Latin or Caribbean countries, while the Institutional Investor is more generous to Asian and European countries than to African ones.
- **Latency:** Another criticism lies in the time taken by the rating agencies to react to new facts (e.g., according to The Economist , "rating agencies may have been too slow to downgrade Japan. Markets have already moved ahead of them").
- **Overreactions:** The IMF criticizes rating agencies claiming that they reacted in panic during the Asian crisis. After they had missed to predict the Asian crisis, they reacted by harshly downgrading countries such as Thailand or South Korea, thus accelerating the flight of capital. In this and other situations, rating agencies gave the impression of overreacting (Figure 1) instead of being a stabilizing force.

Figure 1: Precrisis and postcrisis rating of countries



It appears that the objectivity and reliability of country risk ratings is questionable, mainly because of human intervention and conflicting goals and/or interests.

- ***Negative impact of rating changes:*** It is reported that the hesitation or reluctance of raters to downgrade a country stems from the fact that a downgrade announcement can precipitate a country into crisis. During the Asian crisis, the rating agencies arouse the discontent of the Malaysian Prime Minister, Dr Mahathir bin Mohamad, who condemned them and charged them with rendering the crisis even more acute. "The rating agencies, when we have a need to borrow money, they immediately downgraded us so that it will cost us 15% to borrow money. They stop us completely from borrowing money" (1999)[1]. Along the same line, Reisen and Von Maltzan (1999) claim that such a sharp downgrade impeded "commercial banks to issue letters of credit, forced investors to offload Asian assets to maintain portfolios in investment-grade securities". They argue that rating agencies lagging behind rather than anticipating the state of financial markets reinforce positive expectations and capital inflows when they upgrade countries and intensify outflows of capital and crisis when they downgrade.
- ***Conflicts of interest:*** An even more pointed criticism is that raters, having started charging fees to rated countries, can be suspected of reluctance to downgrade them, because of the possibility of jeopardizing their income sources. This is claimed, for example, by Tom McGuire, an executive vice-president of Moody's, who states that "the pressure from fee-paying issuers for higher ratings must always be in a delicate balance with the agencies' need to retain credibility among investors"[2]. The necessity to please the payers of the ratings, investors as well as issuers, lead to what Robert Grossman, the chief credit officer at the rating agency Fitch, calls "a tendency we do with investors – rating committees, outlooks, meetings, then the press release, all to soften the blow of the rating change"[3]. Studying the rating transitions, Altman and Saunders (1998) notice that a downgrade in the rating of a country is regularly followed by further downward adjustments. The explanation given by Altman and Saunders is that agencies gradually downgrade the rating of a country, since they do not want to hurt the country, which is also their client. Kunczik (2001) note that the IMF (1999) fears the danger that "issuers and intermediaries could be encouraged to engage in *rating shopping* – a process in which the issuer searches for the least expensive and/or least demanding rating".

The problems described above will become more acute as the role of ratings increases. Indeed, the Basel Accord will intensify the pressure on countries to obtain high ratings, potentially leading to a switch from rating shopping to rating fraud. For instance, Pakistan has been forced to pay back $55 million credits to the IMF because of budget falsification, the blame being put on the former Prime Minister Nawaz Sharif, accused of having falsified the budget deficit. Similarly, Ukraine has been proven to have reported misleading data on its reserves in foreign exchanges, attempting to obtain IMF credits. Kunczik (2001) says that "it is only a question of time when firms will specialize in rating advising for sovereigns".

[1] This article appeared in the February 19, 1999 issue *Executive Intelligence Review.*
Interview: Datuk Seri Dr. Mahathir bin Mohamad Malaysian Prime Minister: `We had to decide things for ourselves'. On January 22, 1999, Gail G. Billington of EIR's Asia Desk and Dino de Paoli of the Schiller Institute were given the opportunity to interview Datuk Seri Dr. Mahathir bin Mohamad, Prime Minister of Malaysia.

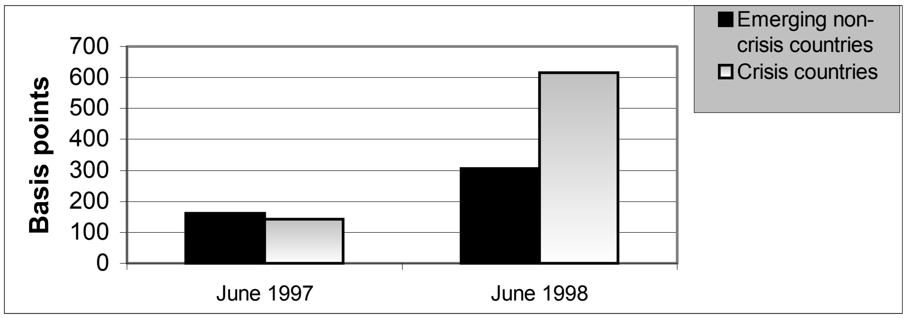
[2] The Economist, July 15, 1995, 62

[3] Euromoney, January 2002, 38, "*Investors turn cool on the rating game*"

1.3 Can yield spreads replace country risk ratings?

To mitigate the problems described above, it is sometimes advocated to use yield spreads instead of sovereign ratings as a proxy for default risk, since large yield spreads correspond to high risk. Yield spreads refer to the difference between sovereign yields and US treasury bill yields of the same maturity. Market yields are less stable, fluctuating daily and sometimes substantially.

Figure 2: Precrisis and postcrisis yield spreads on an aggregate basis



It appears that the use of market spreads rather than country ratings is not more efficient. Indeed, for Asian countries, spreads have substantially widened after the crisis. As exhibited by *Figure 2*, spreads were roughly of the same order of magnitude before the crisis. While spreads of non-crisis countries have widened by less than 100% after the crisis, spreads of crisis countries have more than tripled. Consequently, we conclude that spreads provide about the same information as sovereign ratings do, and are much more volatile. This conclusion can be extended to the Brazilian and the Russian crises. This discussion implies that yield spreads are characterized by a lack of predictive power and cannot be used to obtain a reliable early warning of country insolvency. This latter conclusion is confirmed by Mathieson and Schinasi (1999).

1.4 Recursive versus non-recursive models

The recent literature on country risk ratings contains several studies (Cantor and Packer, 1996, Haque et al., 1998, Monfort and Mulder, 2000) which use multiple regression. The set of independent variables used by Haque et al. (1998) as well as Monfort and Mulder (2000) includes the lagged sovereign ratings of Standard & Poor, or Moody's, or The Institutional Investor. The correlation levels between the ratings of various agencies and the predicted values or ratings obtained using multiple regression models referenced above are remarkably high.

To illustrate the actual meaning and importance of these results, we shall examine the approach taken by Haque et al. (1998). That paper uses as its eight independent variables seven macro-economic variables[4], and the lagged rating, i.e. it includes The Institutional Investor ratings both at times *t* and *t-1*, the former as a dependent variable, and the latter as an independent one. It is important to note that country risk ratings are very stable, as shown by the transition probabilities of the ratings published by Standard & Poor (1999) (see *Table 19* in Appendix) and Nickell et al. (2000).

[4] T-bill rate, GDP growth rate, inflation rate, exports growth rate, ratio current account to GDP, the ratio of external debt to GDP, the ratio of reserves to imports.

The 98% correlation level[5] between The Institutional Investor ratings published respectively in September 1997 and September 1998 confirms the stability property of sovereign ratings. In light of this fact, the excellent correlation levels achieved by utilizing lagged ratings among the independent variables can be attributed to a certain – possibly large – extent to this stability, and may not necessarily give indications about the predictive power of the economic and political variables used as predictors.

Although Cantor and Packer (1996) do not include the lagged ratings in their set of predictors, they create a dummy variable, which is determined by the past ratings issued by Standard and Poor (Claessens and Embrechts, 2002). This dummy variable is defined to be equal to 1 if a country has ever been rated D or SD by Standard & Poor since 1970, and equal to 0 otherwise. Even though their regression *R-square* is above 90%, their results are criticized by Claessens and Embrechts (2002) and Jüttner and McCarthy (2000). Claessens and Embrechts mention that the dates of the explanatory variables are not consistent, e.g. the values of some variables are measured in 1994 or 1995, while that of others are averages for the period 1991-1994 or 1992-1994. On the other hand, Jüttner and McCarthy evaluated the regression model of Cantor and Packer for some other years, concluding that for 1998, it loses its predictive power. A recent paper of Hu et al. (2002) develops a model using ordered probit to estimate country ratings. Their model has an 83% correlation level and relies on economic variables and rating history of countries.

A common feature of the econometric models above is the direct or indirect inclusion of information derived from past Standard & Poor ratings (lagged ratings, rating history) among their independent variables. A major drawback of such rating models is the impossibility of applying them to not-yet-rated countries.

1.5 Objectives and main results

Our discussion in the preceding subsections indicates a need for making country risk ratings more *(i) transparent* and *(ii) consistent*. A third criterion we would like to impose on an ideal country risk rating system is that of *(iii) self-containment*, i.e. its non-reliance on any other past or present country risk ratings. Clearly, this requirement precludes the use of lagged ratings as independent variables. It is important to note that this approach is in marked contrast with that of the current literature (discussed in the previous subsection), which does rely in one form or another on lagged ratings. Finally, a fourth requirement imposed on the model is its *(iv) stability*, i.e. extensibility both to subsequent years and to previously non-rated countries.

The wide acceptance of several of the major rating systems indicates that, while they may not be perfect, they provide the currently best known evaluation of country risk. It is therefore reasonable to base the design of any new rating system on one of the existing ones.

The **central objective** of this paper is to develop a *transparent*, *consistent, self-contained*, and *stable* system, closely approximating the country risk ratings provided by a major rating agency. We have selected the Standard & Poor country risk rating system as a benchmark for the desired system. It is to be expected that, on the one hand, in most cases the ratings of the new system should closely resemble those of Standard and Poor, and on the other hand, in the (hopefully few) cases where the two ratings differ, the objective reasons, which determine the ratings of the proposed model, should be justified by subsequent developments.

[5] 134 countries are considered.

In line with the existing literature, we use in the first part of this paper the technique of multiple linear regression to achieve our objectives. We shall call the proposed system a *non-recursive multiple regression* model of the S&P country rating system. In the second part of the paper, we reanalyze the same problem, using this time a combinatorial-logical technique, in order to derive a set of "logical rating scores" of countries, and show that they turn out to be surprisingly similar to both the S&P ratings and the non-recursive regression scores.

The fact that the traditional linear regression model and the qualitatively different combinatorial-logical model produce almost identical results strongly reinforces the confidence in the results and in the validity of both models. In addition to the main result, we also demonstrate that the new rating systems

- can be successfully applied to the rating of yet non-rated countries,
- are temporally stable,
- are consistent with the ratings of other agencies (Moody and The Institutional Investor).

1.6 Paper structure

The paper is structured as follows. Section 1 describes the data considered and selected for use in this paper. We provide a thorough literature review (see references in Tables 13 and 14) and describe the selection of explanatory variables.

In Section 2, we use the 1998 Standard & Poor country risk ratings to develop a non-recursive multiple regression model for the ratings considered as the dependent variable, regressed on a set of economic and political variables (considered as the predictor variables). To evaluate the accuracy of linear regression predictions, we use the *k*-folding cross-validation technique. We show that the model correlates well not only with the ratings of Standard & Poor (95%), but also with those of Moody's (94.6%) and The Institutional Investor (93.6%).

We also evaluate the stability of the constructed non-recursive regression model in two ways. First, we show the temporal stability of the non-recursive multiple regression model derived from the 1998 dataset by applying it to the 1999 data. Second, we show that the proposed model can successfully predict the ratings of several previously non-rated countries.

In Section 3, we analyze the same problem by using a combinatorial-logical technique, the *logical analysis of data* (LAD) (Hammer, 1986, Crama et al., 1988, Boros et al., 2000), for developing a model that evaluates the creditworthiness of countries. Based exclusively on the S&P ordering of countries by their riskiness, i.e., without making any assumptions about the magnitude of differences between consecutive ratings, we construct a discriminant function, which provides an approximate measure of "relative riskiness" of an arbitrary country i compared with another country j. The discriminant is expressed as a highly nonlinear polynomial in binary variables, which indicate whether the values of relevant economic and political attributes of one country do or do not exceed the values of the corresponding attributes of another country by certain thresholds. The discriminant does not involve any information about past ratings. Further, it is shown how to calculate a numerical measure of country risk (the "logical rating score") in such a way that the differences between the logical rating scores of all pairs of countries provide the best L_2-approximation of the corresponding discriminant values.

Subsection 4.1 introduces the concept of a *pseudo-observation* P_{ij}, associated to a pair of countries i, j, which provides a comparative description of i and j in the form of a multi-dimensional vector, whose components are the differences of the values of those economic and political attributes of countries i and j which were identified in the first part of the study. An additional component of a pseudo-observation is an indicator which takes the value 1 (-1, 0) if the country i in the pseudo-observation has a higher (lower, identical) rating than the country j.

The fundamental idea of this study is that a rating system can be essentially reconstructed from the knowledge of the relative orderings of all pairs of rated countries. In other words, all that matters in a rating is the qualitative order relation between countries, but not a quantitative measure of the magnitude of differences between ratings.

The study focuses on deriving a model of the order relation between countries using the *LAD* methodology which is briefly described in Subsection 4.2. Non-statistical and highly nonlinear, this methodology is not restricted by the satisfaction of the assumptions underlying econometric techniques.

Based on the patterns of the LAD model, a discriminant, $\Delta(P_{ij})$, called *relative preference,* is computed for each pseudo-observation, P_{ij}, The value of $\Delta(P_{ij})$ indicates whether country i should be rated higher or lower than country j. Relying on the assumption that the $\Delta(P_{ij})$ values provide good approximations of the differences of the ratings, the relative preferences are used to derive an approximation of the ratings called *logical rating scores* of countries; these are calculated using multiple linear regression, as described in Subsection 4.3. More precisely, denoting by I the set of countries, the rating of a country k is the regression coefficient, β_k, in the regression model

$$\Delta(P_{ij}) = \beta_i - \beta_j \quad , \text{ for all } i, j \in I, i \neq j .$$

Subsection 4.4 is devoted to a thorough analysis of the results provided by the logical rating score model. It is shown that the correlation between the logical rating scores and the S&P ratings exceeds 95%. Moreover, the correlation between the logical rating scores and the scores provided by the non-recursive linear regression model of Section 3 exceeds 98%.

The robustness of the model was confirmed by jackknife cross-validation. Furthermore, the discriminant derived on the basis of the 1998 S&P ratings was applied to the 1999 data, and the correlation between the resulting logical rating scores and the 1999 S&P ratings is shown to exceed 94%. Since the discriminant does not involve any information about past ratings, it can be used to derive the logical rating scores of previously non-rated countries; this is done in Section 4.4 where it is also shown that the logical rating scores are in agreement with subsequent S&P ratings.

Section 5 presents general conclusions of this study.

2 Data

2.1 Sources

In this paper, we focus on the Standard & Poor country risk ratings. The risk of default is generally defined by Standard & Poor as the probability that a sovereign obligor fails to meet a principal or interest payment on the due date and in full. Standard & Poor's ratings are based on the information provided by the debtors themselves and by other sources considered reliable.

Standard & Poor provides sovereign ratings for local and foreign currency debt. In this paper, we used the foreign currency sovereign ratings. Countries are more vulnerable to foreign currency obligations. An obligor's capacity to repay foreign currency obligations may be lower than its capacity to repay obligations in its local currency, owing to the sovereign government's relatively lower capacity to repay external versus domestic debt. As noted by Cantor and Packer (1996), foreign currency ratings remain the decisive factor in the international bond market. Indeed, foreign currency obligations are more likely to be acquired by international investors than domestic obligations. Foreign currency ratings reflect economic factors, as well as the country intervention risk, i.e. the risk of a country imposing, for example, exchange controls or a debt moratorium, while local currency ratings exclude country intervention risk.

Table 15 in the Appendix lists the different country risk levels or labels used by Standard & Poor, and also provides descriptions associated with these labels. Countries which are assigned a label inferior to BB+ are considered as non-investment grade (speculative) countries. Countries rated CCC+ or lower are regarded as presenting serious default risks. BB indicates the least degree of speculation and CC the highest. Ratings labeled from AA to CCC can be modified by the addition of a plus or minus sign to show relative standing within the major rating categories. We consider such subcategories as separate ratings in our analysis.

We have converted the Standard & Poor rating scale (ranging from AAA to SD) into a numerical scale (ranging from 21 to 0) (see Appendix, *Table 16*) and shall liberally refer to both of them as *S&P ratings*. This type of conversion is commonly used in the literature, see e.g., Bouchet et al. (2003), Estrella (2000), Ferri et al.(2001), Kräussl (2000), Monfort and Mulder (2000), Mulder and Perelli (2001), Hu et al. (2002), Sy [2003]. Moreover, Bloomberg, a major provider of financial data services, developed a standard cardinal scale for comparing Moody's, S&P and Fitch-BCA ratings (Kaminsky and Schmukler, 2002); in this scales, a higher numerical value denotes a higher probability of default.

In *Table 16* of the Appendix, we display Standard & Poor's foreign currency sovereign ratings of 69 countries published at the end of December 1998. Standard & Poor rates a limited number of countries, with a special focus (at least in the past) on the industrial ones. However, in the last decade, the number of Asian, Latin American and Eastern European economies rated by Standard & Poor has significantly increased. We refer the reader to Hu et al. (2002) for the evolution of the number of countries rated by Standard & Poor.

As mentioned above, country risk ratings encompass economic, financial and political aspects. The statistical data of the economic and financial variables considered in this paper come from the International Monetary Fund (World Economic Outlook database), from the World Bank (World Development Indicators database) while those about the ratio of debt to gross domestic product come from Moody's publications. Values of political variables are provided by Kaufmann et al. in two papers (1999a,b) that are joint products of the Macroeconomics and Growth, Development Research Group and Governance, Regulation and Finance Institutes which are affiliated with the World Bank. Before describing the relevance of the selected variables, we discuss in Section 2.2 the selection method used.

2.2 Variable selection criteria

As underlined by Bilson et al. (2001), the selection of variables lends itself to criticism due to the subjectivity and arbitrariness involved in this process. In this paper, the selection of relevant variables is based on three criteria.

The first criterion is the significance of variables for estimating a country's creditworthiness. We have performed an extensive literature review which played an important role in defining the set of candidate variables for inclusion in our model. *Tables 13* and *14* list variables that have been considered in the existing literature on country risk.

The second criterion is the availability of complete and reliable statistics. We want to avoid difficulties related to missing data that could reduce the statistical significance and the scope of our analysis. For instance, according to recent information received from The World Bank[6], their research concentrates on developing economies and they have data on the debt of 137 countries to whom they loan funds and who report their external debt to The World Bank. Since high income countries do not receive World Bank funds, they do not report their debt numbers to The World Bank. Such situations have significantly complicated the process of compiling complete debt statistics. Hu et al. (2002) also report the problem of data availability.

The third criterion is the uniformity of data across countries. We have considered, for example, incorporating the unemployment rate statistics disclosed by the World Bank. However, the World Bank underlines that unemployment is analyzed and compiled according to definitions which differ from country to country.

It is worth noting that in addition to the variables listed in *Tables 13* and *14* (see Appendix), Haque et al. (1996), Cantor and Packer (1996), Larrain et al. (1997), Monfort and Mulder (2000) and Hu et al. (2002) use a dummy variable that represents the historical solvency of a country. Haque et al. (1996) use the lagged rating at time (t-1) as an independent variable in their regression model. Monfort and Mulder (2000) claim that membership in the OECD is likely to be a significant indicator for country risk ratings. The same authors emphasize also the importance of the location of countries, by adding to their set of independent variables two dummy variables to characterize the country's location in Asia or in Latin America. Hu et al. (2002) also use regional dummy variables.

2.3 Selected variables

Based on the criteria of relevance, availability and uniformity described above, we have decided to incorporate the following variables[7] in our model:

Gross domestic product per capita[8] *(GDPc):* the *GDP* is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. The *GDP* is an indication of the capacity of the government to solve a balance-of-payments crisis without having to default on external debt. The larger the *GDP*, the wider the potential tax base and thus the higher the ability of the government to fulfill its external obligations. The *GDPc* is a measure of the relative wealth of a country and its level of development. The gross domestic product (*GDP*) is converted to international dollars using purchasing power parity rates. The international dollar has the same purchasing power over *GDP* as the U.S. dollar has in the United States.

Inflation rate (IR): the inflation rate is the change in the national price level between two periods. The inflation rate used in our study is based on the consumer price index and is the annual percentage change in the cost to the average consumer of acquiring a fixed basket of goods and services. High inflation rates indicate structural problems in the country's finances and may lead to sovereign economic crises, as governments hike interest rates sharply in order to strengthen their countries' currencies. Should a country be unable or unwilling to pay the current budgetary

[6] Anat Lewin, Private Communication, Development Data Group, The World Bank, June 06, 2001

[7] Acronyms in parentheses following the name of variables are used in tables and appendices for referring to variables.

[8] Calculated on the basis of purchasing power parity in international dollars.

expenses, it must resort to inflationary money financing. High inflation rate results in a substantial consumers' purchasing power reduction and increases political discontent.

Trade balance (TB): trade balance is the balance of trade in goods expressed as a percentage of *GDP* (purchasing power parity-PPP). This is the difference in value between a country's total imports and exports (including information of oil and non oil exports, consumer goods, capital goods) measured in current U.S. dollars divided by the value of *GDP* converted into international dollars using purchasing power parity rates.

Exports' growth rate (EGR): annual growth rate of exports of goods and services based on constant local currency. Exports of goods and services represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude labor and property income as well as transfer payments. Countries having a high export growth rate are expected to be more creditworthy. Indeed, exports are the primary source of foreign currency inflows and therefore have a significant influence on the capacity of the country to finance imports and service debt obligations.

International reserves (RES): this variable refers to gross international reserves, expressed in terms of the number of months for which the existing reserves can cover the cost of imports of goods and services. It gives an indication of the short-term capacity of an economy to meet its imports obligations. The higher the value of *RES*, the lower the risk of default and the higher the creditworthiness.

Fiscal balance (FB): fiscal balance is approximated by the ratio of central government financial balance (surplus or deficit) to *GDP*. The central government's balance represents the yearly fiscal balance. Fiscal balances and debt stocks of governments are crucial indicators when analyzing sovereign risk. The ability of governments to extract revenues from taxpayers and users of services is a key factor that helps to determine whether governments will be able to make full and timely payments of interest and principal on outstanding debt.

Debt to GDP (DGDP): here debt refers to the general government debt. The general government debt as defined by the IMF (2001) includes "the consolidated budgets of the central, state/regional, and local governments, along with the social security system and other extra-budgetary funds engaged in noncommercial activities. Excluded are lending and refinancing and the assets/liabilities of commercial state-owned or guaranteed enterprises, except for any net financial transfers made as subsidies to these enterprises". This balance, i.e., the difference between total revenues and total expenditures, determines the net borrowing requirement of general government, which can be met only by running down financial assets or borrowing net new resources from the public and, thereby, adding to debt.

We have considered incorporating the unemployment rate and the ratio of the current account balance to GDP. While the latter turned out to be redundant with trade as a percentage of GDP, the former has been excluded from consideration due to the lack of consistency in its definition. As noted by the World Bank, the treatment reserved to temporarily laid off workers, to those looking for their first job, and the criteria referred to for being considered as unemployed, differ significantly between countries.

For political variables, it is very difficult to find reliable and complete data. In our model, we have considered the six variables provided by Kaufmann et al. (1999a). These six variables are: political stability and violence, voice and accountability, government effectiveness, regulatory burden, corruption , rule of law. These variables are viewed as capturing the fundamentals of the governance concept defined as "the traditions and institutions by which authority in a country is exercised" (Kaufmann et al.,1999a).

As emphasized by Kaufmann et al. (1999, a and b), political stability and voice and accountability both refer to the process by which governments are elected, monitored and replaced. Government effectiveness and regulatory burden reflect the capacity of the government to adopt sound policies. Corruption and rule of law are proxies for the "respect of citizens and institutions for the rules which govern their interactions". In order to avoid or at least limit redundancies in our model, we select only one variable for each dimension of governance. We have selected:

Political stability (PS),
Government effectiveness (GE), and
Corruption (COR).

The higher the values of these variables, the less likely the country is to default[9]. The variables are defined on a (-3.5, 3.5) interval and are based on estimations provided by polls of experts and cross-country surveys.

The variables we have described so far have been considered previously in the literature and are available in the form used in our study (as ratios or as growth rates). We have also decided to construct a new variable (*ER*) and to add a variable (financial depth and efficiency) which, to the best of our knowledge, has not been used before in country rating studies. Here are the descriptions of these two variables:

Exchange rate (ER): is defined as the ratio of the current value of the exchange rate to the moving average of the real effective exchange rate[10] over five years (1994 to 1998). While the exchange rate has been used in previous country rating studies, we consider the ratio introduced here to be more significant, since it indicates the dynamics of changes in the exchange rate, by specifying whether the trend is up ($ER>1$) or down ($ER<1$).

Financial depth and efficiency (FDE): is represented by the ratio of the domestic credit provided by the banking sector to the *GDP*. Households accumulate claims on financial institutions that, acting as intermediaries, pass funds to final users. Correlated to the development of the economy, the indirect lending by savers to investors becomes more efficient and gradually increases assets relative to the *GDP*. Viewed from this perspective, the ratio of domestic credit to the *GDP* reflects the financial depth and efficiency of the country's financial system. More specifically, this variable is used to measure the growth of the banking system since it reflects the extent to which savings are financial. To our knowledge, the financial depth and efficiency variable has not been considered previously in the evaluation of country risk ratings.

2.4 Dataset content

In summary, on the basis of the considerations described above, we have constructed a dataset involving nine economic/financial variables:

- gross domestic product per capita, inflation rate, trade balance, international reserves, fiscal balance, exports growth rate, debt to *GDP*, financial depth and

[9] The higher the value of the corruption variable, the less corrupted the considered country is perceived to be. This variable can therefore be called "corruption quality".

[10] Real effective exchange rate is the nominal effective exchange rate (a measure of the value of a currency against a weighted average of several foreign currencies) divided by a price deflator or index of costs.

efficiency, and exchange rate (we have used the values taken by these variables at the end of 1998);

and three political variables:

- political stability, government effectiveness and corruption level.

We have compiled the values of these twelve variables for the sixty-nine countries considered: 24 industrialized countries, 11 Eastern European countries, 8 Asian countries, 10 Middle Eastern countries, 15 Latin American countries and South Africa. We use the Standard & Poor country risk ratings for these countries at the end of December of 1998.

3 A Statistical Model

3.1 Non-recursive multiple regression model

3.1.1 The model, results, and consistency with S&P

In order to derive a non-recursive model of Standard & Poor's ratings, we shall fit the regression equation:

$$Y = \alpha + \sum_{i=1}^{M} \beta_i * X_i + \varepsilon \quad , \tag{1.1}$$

where the dependent variable Y is the country risk rating given by Standard & Poor at the end of December 1998 (or more precisely a numerical representation of Standard & Poor's ratings), the independent variables X_i are the economic and political variables described in Section 2.1, and ε is the error term. In view of the desired non-recursiveness of the model, the independent variables do not include directly or indirectly ratings of previous years. Results given in this section have been obtained using the SPSS statistical package.

The proposed model exhibits an excellent fit, with the coefficient of multiple determination *R-square* being 91.2%, and the adjusted *R-square* 89.3%. The multiple correlation level between the observed values (i.e. the Standard & Poor ratings) and the predicted ones (i.e. the ratings given by the non-recursive regression model) is equal to 95.5%. The later are given in Appendix (Table 17, Column 2).

Table 1 below details how the regression equation accounts for the variability in the response variable, the last column giving the statistical level (*1-p*) at which the model is significant.

Table 1: Analysis of variance (ANOVA)

	Sum of Squares	Degrees of freedom	Mean Square	F-statistic	*p*-value
Regression	1609.546	12	134.129	48.458	0.000
Residual	155.005	56	2.768		
Total	1764.551	68			

Table 2 presents the regular and the standardized regression coefficients (i.e., those corresponding to the model fitted to standardized data). The last column in *Table 2* indicates whether the corresponding independent variable is statistically significant (at the confidence level of *1-p*).

Table 2: Regression results

Variables	Unstandardized coefficients	Standard error	Standardized coefficients (Beta)	t-statistic	p-value
Intercept	8.769	.860		10.195	0.000
FDE	1.693E-02	0.007	0.148	2.513	0.015
RES	0.116	0.101	0.055	1.148	0.256
IR	-2.831E-02	0.018	-0.080	-1.557	0.125
TB	-1.192E-02	0.006	-0.094	-1.960	0.055
EGR	-1,218E-02	0.031	-0.017	-0.396	0.694
GDPC	3.081E-04	0.000	0.499	5.294	0.000
ER	-1.968E-02	0.011	-0.079	-1.768	0.083
FB	0.120	0.086	0.078	1.393	0.169
DGDP	-1.610	0.795	-0.091	-2.026	0.047
PS	1.378	0.533	0,197	2.584	0.012
GEF	1.977	0.920	0.316	2.149	0.036
COR	-0.605	0.842	-0.111	-0.718	0.476

At the 5% significance level, it appears that five independent variables are statistically significant. These are: financial and depth efficiency (*FDE*), gross domestic product per capita (*GDPc*), ratio debt to gross domestic product (*DGDP*), political stability (*PS*) and government efficiency (*GE*).

The regression results described above indicate that the non-recursive regression model has an excellent fit with the data. However, the excellence of the fit does not automatically guarantee the predictive power of the model, if the model violates some of the critical assumptions of multiple regression theory, as is the case with proposed model. Indeed,

- there is a strong correlation between some of the variables considered, e.g. between the political variables, especially government efficiency and corruption, possibly leading to difficulties related to multicollinearity, and the ill-conditioned nature of the resulting matrix;
- the predictors are not normally distributed;
- if too many predictor variables are used relatively to the number of observations, fitting multiple regression can lead to overfitting, and the estimates of the regression line can be unstable and the results may not be reproducible; the number of variables used is generally recommended to be no more than 5 to 10% of the number of observations, which is clearly not the case of this study that involves 69 observations and 12 variables.

In view of these issues, it is surprising that the cross-validation results presented in the next section provide a strong confirmation of the predictive power of the non-recursive regression model presented above.

3.1.2 Cross-validation

To validate the predictive power of the non-recursive regression model, we use a resampling technique known as cross-validation, and more specifically, a popular variant of it called *k*-folding (e.g., Shao, 1993, Shao and Tu, 1995, Efron, 1982, Hurvich and Tsai, 1989, Hjorth 1994, Breiman and Spector, 1992). In *k*-folding, observations are divided into *k* subsets of approximately equal size. The regression model is trained *k* times, each time leaving out from training one of the *k* subsets, and using the omitted subset to test the regression-predicted country risk rating. In this paper, based on the relatively small size of the sample, we have selected *k* to be 10, and partitioned the sample into 10 groups of 6 or 7 countries each. The groups were selected using stratified random sampling, i.e. assuring that each group contains about the same number of investment-, speculative- and default-grade countries (see S&P's classification, *Table 15*).

In Appendix (Table 17, Column 3) we present the in-the-sample predictions of the non-recursive multiple regression model obtained in the preceding Section, and the out-of-the-sample predictions obtained using the 10-fold cross-validation. The major results are the following:

- the correlation between the in-the sample and the out-of-the sample predictions is **99.1**%,
- the correlation between the Standard and Poor ratings and the out-of-the sample predictions is **95.6**%.

The very high correlation levels demonstrate clearly that the impressive results of Section 3.1.1 are not due to chance or overfitting.

3.1.3 Rating discrepancies between S&P and the proposed model

In this section, we shall identify the few countries for which the predictions of the non-recursive regression model disagree with the Standard & Poor ratings. In order to accomplish this, we shall construct confidence intervals for our predicted ratings[11].

Let us introduce some notations. Let n and p refer to the number of observations and predictors, respectively. The expression $t(1-\alpha/2, n-p)$ refers to the Student test with $(n-p)$ degrees of freedom, and with upper and lower tail areas of $\alpha/2$. Let X_j be the p-dimensional vector of the values taken by the observation Y_j on the p predictors, while $X_p^{'}$ be the transposed of X_j. Let the expression $(X'X)^{-1}$ refer to the variance-covariance matrix, i.e. the inverse of the $[p\times p]$-dimensional matrix $(X'X)$. Denoting by MSE the mean square of errors in the regression, the estimated variance $s^2[\hat{Y}_j]$ of the predicted rating is:

$$s^2[\hat{Y}_j] = MSE*[X_j^{'}(X'X)^{-1}X_j] \quad , \tag{1.2}$$

while the $(1-\alpha)$-confidence interval for the predicted rating $\hat{Y}_j$ is:

$$\{\hat{Y}_j - t(1-\alpha/2, n-p)*s[\hat{Y}_j], \hat{Y}_j + t(1-\alpha/2, n-p)*s[\hat{Y}_j]\} \tag{1.3}$$

[11] All formulae given in this section as well as those in Section 3.2.3 are from Neter et al. (1996)

We say that there is a discrepancy between the Standard & Poor rating R_j^{SP} of a country j and ours, if the Standard & Poor rating is not in the confidence interval, i.e.:

$$R_j^{SP} \notin \{\hat{Y}_j - t(1-\alpha/2, n-p) * s[\hat{Y}_j], \hat{Y}_j + t(1-\alpha/2, n-p) * s[\hat{Y}_j]\} \text{ for } \alpha = 0.1 \qquad (1.4)$$

Taking α equal 5%, this formula identifies four discrepancies. Three countries (Iceland, Pakistan and Argentina) are rated higher by the non-recursive regression model than by Standard & Poor, while Columbia is rated higher by Standard & Poor. It is remarkable that subsequently the Standard & Poor ratings for two of these four countries (Columbia and Pakistan) have been modified in the direction suggested by the regression model. More precisely, Columbia has been downgraded by Standard & Poor twice, moving from BBB- in December 1998 to BB+ in September 1999, and then to BB in March 2000. After being downgraded in January 1999 (SD), Pakistan was upgraded to B- in December 1999. On the other side, Iceland's rating has remained unchanged, and Argentina's rating has endured significant downgrade, which started however only in November 2000.

3.1.4 Are political variables necessary?

In this section, we test the predictive power of the non-recursive regression model, from which the three political variables are omitted. The *R-square* as well as the adjusted *R-square* of this model are equal to 88.6 % and 86.9 % respectively. These values are lower than the corresponding values for the original non-recursive regression model, indicating a loss in predictive power resulting from the omission of the three political variables. The predicted ratings are given in Appendix (Table 17, Column 4).

Table 3: Regression coefficients

Variables	Unstandardized coefficients	Standard error	Standardized coefficients	t-statistic	p-value
Intercept	1.463	2.631		0.556	0.58
FDE	2.20E-02	0.007	0.192	3.058	0.003
RES	0.194	0.107	0.091	1.818	0.074
IR	1.22E-02	0.034	0.034	0.359	0.721
TB	-5.85E-03	0.006	-0.046	-0.901	0.371
EGR	2.75E-02	0.032	0.039	0.849	0.399
GDPC	4.38E-04	0	0.709	10.776	0
ER	6.933	2.737	0.23	2.533	0.014
FB	0.244	0.083	0.158	2.933	0.005
DGDP	-1.815	0.853	-0.102	-2.129	0.037

It appears that five of the independent variables are statistically significant at a 95% level. These variables are financial and depth efficiency (*FDE*), gross domestic product per capita (*GDPc*), debt to gross domestic product ratio (*DGDP*), exchange rate (*ER*) and fiscal balance (*FB*). The correlation coefficient between the predicted ratings and those of Standard & Poor is equal to 94.14 % and is lower than in the original model. Moreover, the inferior fit of this model results in wider confidence intervals as compared to the original model.

The discrepancies between Standard & Poor's predictions and ours involve four countries (Russia, Pakistan, South Korea and Iceland), all being underrated by Standard & Poor. The ratings of three of these countries (Russia, Pakistan, South Korea) have been modified since, in the direction suggested by our model, while the rating of Iceland has remained unchanged. The evolution of ratings for Pakistan has already been described in Section 3.1.3. South Korea has both been upgraded three times, moving from BB+ in December 1998 to BBB+ in November 2001. Russia has first been downgraded to SD in January 1999, before being upgraded three times and being rated B+ in December 2001.

In conclusion, the model which omits political variables appears to be somewhat less closely related to the S&P model which it is supposed to reflect, but on the other end this apparent weakening is not sufficiently clear to allow us to draw any definite conclusions. It should be added here that the economic variables are easier to obtain than the political ones, which are published less frequently.

3.2 Stability of the non-recursive regression model

3.2.1 Consistency with ratings of other agencies

In addition to analyzing the correlation level between Standard & Poor's ratings and those of the proposed non-recursive model, the latter has to be compared with the ratings of other agencies, e.g. Moody's and The Institutional Investor. We present below the results of these comparisons, based on Moody's and The Institutional Investor ratings issued at the end of December 1998 and in March 1999 respectively. We shall start by presenting a brief description of the rating systems of Moody's and of The Institutional Investor.

Moody's sovereign ratings are defined, as "a measure of the ability and willingness of the country's central bank to make available foreign currency to service debt, including that of central government itself" (Moody's, 1995). Similarly to Standard & Poor, Moody's uses a nominal rating scale (Table 18 in Appendix), which contains the same number of categories as Standard & Poor's ratings. A large proportion of countries receive the same rating from Moody's and Standard & Poor, and when they are different, the difference is usually not more than one notch.

The Institutional Investor country risk ratings were first compiled in 1979, and are published now regularly, in March and September of every year, for an increasing number of countries, which reached 145 in 2000. The Institutional Investor ratings are numerical, ranging from 0 to 100, with 100 corresponding to the lowest chance of default. The Institutional Investor relies on evaluations of the creditworthiness of the countries to be rated, provided by economists and international banks, each respondent using their own criteria. Responses are aggregated by The Institutional Investor, greater weights being given to responses from institutions with higher worldwide exposure.

The correlation levels between the ratings given by the non-recursive multiple regression model and those given by Standard and Poor, Moody's and The Institutional Investor are reported in Table 4.

Table 4: Correlation between the non-recursive model and other ratings

	Non-recursive regression	Standard & Poor	Moody's	The Institutional Investor
Non-recursive regression	100%	95.05%	94.56 %	93.60 %
Standard & Poor		100%	98.01%	96.18%
Moody's			100%	96.31%
The Institutional Investor				100%

It can be seen that the very high correlation levels between the ratings given by the non-recursive regression model and those given by the three rating agencies underline the relevance of the proposed model.

3.2.2 Temporal stability of the non-recursive regression model

3.2.2.1 1999 data

In order to supplement the indications of stability of the non-recursive regression model provided by cross-validation, we shall test its temporal stability by extending the analysis to the data of the following year (1999). Using as input the same economic and political variables described in Section 2.3, we shall use the 1999 data in two ways. First, we shall use the model derived from the 1998 data to check the quality of its predictions when applied to the 1999 data. Second, we shall derive a new model based on the 1999 data and check the goodness of its fit.

In this experiment, we have used 1999 data with two exceptions. First, in the 1999 data, 16 out of the 828 variable values (1.9%) are missing; for these missing values, we have substituted their corresponding values taken in the previous year. Second, the political variables, reflecting the perceptions of governance quality by a large number of survey respondents in industrial and developing countries, as well as in non-governmental organizations, are not necessarily compiled and updated on a yearly basis. The indices of Kaufmann et al. (1999a, 1999b and 2002) have been published twice, referring respectively to data of 1998 and 2000-2001, but were not compiled for 1999. The high degree of stability of political variable indices is reflected in the fact that the correlation between the 1998 and the 2000-2001 indices varies between 95% and 98%, depending on the variable. In our calculations, we have approximated the values of the three political variables appearing in our model (corruption, government efficiency and political stability) for 1999, by averaging their values for 1998 and 2000-2001.

3.2.2.2 Applying the 1998 model to the 1999 data

The purpose of this section is to test the applicability of the non-recursive regression model built on the 1998 data, for predicting the 1999 country risk ratings. More precisely, we substitute the 1999 data into the regression model built on the 1998 data, and compare the results obtained in this way with Standard & Poor's 1999 ratings. The new predicted country risk ratings are given in Appendix (Table 17, Column 5).

The most important result of this experiment is the very high level of correlation (94.74%) between the predicted ratings and the 1999 Standard & Poor ratings, confirming the consistency and the temporal stability of the non-recursive regression model.

In order to identify the discrepancies, we have to recalculate the prediction confidence intervals for the new, 1999 observations. Since these observations have not been used in deriving the regression coefficients, formulae (1.2), (1.3), and (1.4) can no longer be used for constructing the confidence intervals. The variance $s^2[pred]$ should now be computed as follows:

$$s^2[pred] = MSE * [1 + X_j^{'}(X'X)^{-1} X_j] \quad , \tag{1.5}$$

while the $(1-\alpha)$ confidence interval for $\hat{Y}_{j,n}$ will be given by:

$$\{\hat{Y}_{j,n} - t(1-\alpha/2, n-p) * s[pred], \hat{Y}_{j,n} + t(1-\alpha/2, n-p) * s[pred]\} \tag{1.6}$$

We say that there is a discrepancy between the Standard & Poor rating R_j^{SP} and the non-recursive regression model, if :

$$R_j^{SP} \notin \{\hat{Y}_{j,n} - t(1-\alpha/2, n-p) * s[pred], \hat{Y}_{j,n} + t(1-\alpha/2, n-p) * s[pred]\} \text{ for } \alpha = 0.1 \tag{1.7}$$

The results show that the only three discrepancies between our ratings and those of Standard & Poor concern Argentina, Iceland and Russia, all of these countries being underrated by Standard & Poor. The cases of Argentina and Iceland have already been discussed in Section 3.1.3. As far as Russia is concerned, it was first downgraded to SD in January 1999, but upgraded afterwards to B- in December 2000, and to B in June 2001, thus confirming our prediction.

3.2.2.3 New model for 1999

To obtain a new non-recursive regression model for 1999, we proceed in the same way as for 1998, using as input the political and economic variables described above and the sovereign risk ratings published by Standard & Poor at the end of December 1999. As a result, we obtain new regression coefficients that constitute the 1999 model. The 10-folding cross-validation tests performed showed a correlation level of 98.63% between the in-the-sample and out-of-the-sample predicted ratings, thus validating the 1999 model.

The final non-recursive regression model for 1999 is given in Table 5 and the predicted ratings are given in Appendix (Table 17, Column 6).

It is important to emphasize that the correlation between the 1999 Standard & Poor ratings and those predicted by the 1999 non-recursive regression model is 96.4% (even exceeding that of 1998). The only discrepancies between the two models, determined using (1.4), concern Iceland and Argentina, which appear to be underrated by Standard & Poor.

Table 5: Regression results

Variables	Unstandardized coefficients	Standard error	Standardized coefficients	t-statistic	p-value
Intercept	9.97	2.065		4.829	0
Intercept	9.97	2.065		4.829	0
FDE	0.007446	0.006	0.065	1.312	0.195
FDE	0.007446	0.006	0.065	1.312	0.195
RES	-0.002063	0.093	-0.001	-0.022	0.982
RES	-0.002063	0.093	-0.001	-0.022	0.982
IR	-0.0875	0.025	-0.236	-3.454	0.001
IR	-0.0875	0.025	-0.236	-3.454	0.001
TB	-0.005184	0.005	-0.046	-1.003	0.32
TB	-0.005184	0.005	-0.046	-1.003	0.32
EGR	-0.023	0.023	-0.038	-0.989	0.327
EGR	-0.023	0.023	-0.038	-0.989	0.327
GDPC	0.0002413	0	0.406	4.894	0
GDPC	0.0002413	0	0.406	4.894	0

ER	-0.798	2.274	-0.025	-0.351	0.727
FB	-0.05838	0.071	-0.041	-0.819	0.416
FB	-0.05838	0.071	-0.041	-0.819	0.416
DGDP	-0.856	0.727	-0.048	-1.178	0.244
DGDP	-0.856	0.727	-0.048	-1.178	0.244
PS	0.786	0.569	0.108	1.382	0.173
PS	0.786	0.569	0.108	1.382	0.173
GEF	2.722	0.946	0.422	2.878	0.006
GEF	2.722	0.946	0.422	2.878	0.006
COR	-0.04607	0.792	-0.008	-0.058	0.954
COR	-0.04607	0.792	-0.008	-0.058	0.954

3.2.3 Rating previously non-rated countries

In this section, we test the prediction power of our model on a set of countries which were not used for constructing our model. We have obtained data for four such countries (Ecuador, Guatemala, Jamaica and Papua New Guinea), the ratings of which by Standard & Poor started after December 1998. The ratings of these countries using the 1998 non-recursive regression model (described in Section 3) with 1998 and 1999 data are presented in Table 6.

Table 6: Predicted ratings for previously non-rated countries

	Time of the first S&P rating	S&P rating	S&P linear extension	1998 regression model rating using:	
				1998 data	1999 data
Ecuador	07/2000	(SD)	(0)	6.03	5.69
Guatemala	10/2001	(BB)	(10)	8.57	8.14
Jamaica	11/1999	(B)	(7)	6.00	7.16
Papua New Guinea	01/1999	(B+)	(8)	6.93	6.88

Comparing the ratings predicted by the non-recursive regression model with those given by Standard & Poor, it can be seen that the model has a significant predictive power, even when applied to countries not used in constructing the model. Indeed, three of the four countries above (Papua New Guinea, Jamaica and Guatemala) have their first Standard & Poor ratings within the 95% confidence intervals of the predicted ratings. The rating of Ecuador is even more interesting. While the first Standard & Poor rating of that country (SD) is not in the 95% confidence interval of the predicted value, that rating (given in July 2000) was revised after only one month (in August 2000) to B, which falls within the 95% confidence interval of our prediction.

4 A Combinatorial Model

4.1 Pairwise country comparisons: Pseudo-observations

Let us associate to every country $i \in I = \{1,\ldots,69\}$ considered in this study, the 13-dimensional vector C_i, whose first component is the country risk rating given by Standard and Poor, while the remaining 12 components specify the values of the nine economic/financial and of the three political variables.

In this study, instead of considering countries independently of each other, we shall consider *pairs of countries*. For this purpose, we shall construct for every pair of countries $i, j \in I$, a *pseudo-observation* P_{ij} , which shall provide in a way specified below a comparative description of the two countries.

The pseudo-observations are represented as 13-dimensional vectors. The first component is an indicator which takes the value 1 if the country i in the pseudo-observation P_{ij} has a higher rating (i.e., lower risk) than the country j, takes the value –1 if the country j has a higher rating than the country i, and takes the value 0 if both countries have the same rating. The other components k , $k = 2,\ldots,13$ of the pseudo-observation $P_{ij}[k]$ are obtained simply by taking the differences of the corresponding components:

$$P_{ij}[k] = C_i[k] - C_i[k], k = 2,\ldots,13 \tag{1.8}$$

One of the advantages of this transformation is that it allows us to avoid the problems posed by the fact that the original dataset contains only a small number ($| I |$) of observations. The transformation (1.8) provides a substantially larger dataset, which contains $| I |*(| I | -- 1)$ pseudo-observations. It will be seen that the potential problems created by the non-independence of pseudo-observations can be overcome by robust data analysis techniques.

The fundamental idea of this study is that a rating system can be essentially reconstructed from the knowledge of the relative standings of all pairs of rated countries. In other words, all that matters in a rating is the order relation between countries. Therefore, this study will focus on inducing a model for the order relation between countries.

In order to illustrate the construction of pseudo-observations, let us consider as an example the case of Japan and Canada.

Table 7 reports the values taken by the twelve economic/financial and political variables, as well as the rating given by Standard & Poor to these countries at the end of December 1998.

Table 7: Examples of country observations

	S&P RATING	*FDE*	*RES*	*IR*	*TB*	*EGR*	*GDPc*	*ER*	*FB*	*DGDP*	*PS*	*GE*	*COR*
C_{Japan}	AAA	138.44	5.168	.65	21.7471	-2.54	24314.2	0.839	-7.7	0.47	1.153	0.839	0.724
C_{Canada}	AA+	94.69	1.01964	0.99	55.9177	8.79	24855.7	0.939	0.9	0.5	1.027	1.717	2.055

In Table 8 below, we construct the pseudo-observation $P_{Japan,Canada}$ from the country observations C_{Japan} and C_{Canada} given in Table 7. Since Sweden and Australia are rated respectively AA+ and AA by Standard & Poor at the end of December 1998 and the rating AA+ is superior to the rating AA, the first component of the pseudo-observation vector takes the value 1. Clearly the pseudo-observation $P_{Canada,Japan}$ is anti-symmetric to $P_{Japan,Canada}$, as shown in Table 8.

Table 8: Examples of pseudo-observations

	Indicator	*FDE*	*RES*	*IR*	*TB*	*EGR*	*GDPc*	*ER*	*FB*	*DGDP*	*PS*	*GE*	*COR*
$P_{Japan,Canada}$	1	43.75	4.15	-0.34	-34.17	-11.33	-541.5	-0.1	-8.6	-0.03	0.126	-0.878	-1.331
$P_{Canada,Japan}$	-1	-43.75	-4.15	0.34	34.17	11.33	541.5	0.1	8.6	0.03	-0.126	0.878	1.331

4.2 Logical Analysis of Data (LAD) – An Overview

The *logical analysis of data* is a combinatorics-, optimization-, and Boolean logic-based methodology for analyzing archives of observations (Boros et al., 1997). Initially created for the classification of binary data (Hammer 1986), LAD was later extended (Boros et al., 1997) to datasets having binary (positive and negative) observations, depending on numerical variables. LAD distinguishes itself from other pattern recognition methods and data mining algorithms by its capacity to discover a minimal (irredundant) set of variables along with a collection of patterns built on them, in order to explain the positive or negative nature of the observations in an dataset (Boros et al., 2000).

Each observation is represented by an *(n+1)*-dimensional vector, the first component of which is the classification of the observation, the *n* other variables being the inputs. The classification is binary (0 or 1), i.e., it specifies the positive (1) or negative (0) nature of the observation.

The purpose of LAD is to discover a "function" *f* depending on the *n* input variables, or an approximation of *f* , allowing the correct discrimination between positive and negative observations. The approximation of *f* is constructed as a weighed sum of patterns. *Positive (negative) patterns* are combinatorial rules which impose upper and lower bounds on the values of a subset of input variables, such that:

- a sufficiently high proportion of the positive (negative) observations in the dataset satisfy the conditions imposed by the pattern, and
- a sufficiently high proportion of the negative (positive) observations violate at least one of the conditions of the pattern.

The conditions defining a pattern specify that the values of some of the variables are "large" or are "small"; more precisely, these conditions require these values to be above or below certain specified levels, called *cutpoints*. By associating an indicator variable to each cutpoint, the dataset is "binarized", i.e., each original numerical variable is replaced by several binary ones.

The following terminology will be useful. The *degree* of a pattern is the number of variables the values of which are bounded in the definition of the pattern. The *prevalence* of a positive (negative) pattern is the proportion of positive (negative) observations covered by it. The *homogeneity* of a positive (negative) pattern is the proportion of positive (negative) observations among those covered by it. Patterns of low degree, high prevalence and high homogeneity have been shown to be the most efficient in LAD applications, e.g., (Boros et al., 2000).

The first step in applying LAD to the dataset is to generate the *pandect*, i.e., the collection of all patterns in an dataset. The number of patterns contained in the pandect of a dataset of such dimensions can be exponentially large, in the order of hundreds of thousands, possibly millions. Because of the enormous redundancy in this set, we shall impose a number of limitations on the set of patterns to be generated, by restricting their degrees (to low values), their prevalences (to high values), and their homogeneities (to high values); these bounds are known as LAD *control parameters*. It should be added that the quality of patterns satisfying these conditions is usually much higher than that of patterns having high degrees, or low prevalences, or low homogeneities. Several algorithms have been developed for the efficient generation of substantial subsets of the pandect corresponding to reasonable values of the control parameters (Alexe and Hammer 2001; Hammer, Kogan et al. 2001; Alexe, Alexe et al. 2002; Alexe and Hammer 2002).

The substantial redundancy among the patterns of the pandect makes necessary the extraction of (usually small) subsets of positive and negative patterns, sufficient for classifying the observations in the dataset. Such collections of positive and negative patterns are called *models.* A model is supposed to contain positive (negative) patterns « covering » (i.e., whose conditions are satisfied by) each of the positive (negative) observations in the dataset. Furthermore, good models tend to minimize the number of points in the dataset covered by both positive and negative patterns in the model.

The way a LAD model can be used for classification is the following. An observation (whether it is contained or not in the given dataset) which satisfies the conditions of some of the positive (negative) patterns in the model, but which does not satisfy the conditions of any of the negative (positive) patterns in the model, is classified as *positive (negative)*. An observation satisfying both positive and negative patterns in the model is classified with the help of a discriminant which assigns specific weights to the patterns in the model (Boros et al., 2000). More precisely, if p and q represent the number of positive and negative patterns in a model, and if h and k represent the numbers of positive, respectively negative patterns in the model covering a new observation θ, then the value of the discriminant $\Delta(\theta)$ is simply

$$\Delta(\theta) = h/p - k/q, \tag{1.9}$$

and the corresponding classification is determined by the sign of this expression. Finally, an observation for which $\Delta(\theta) = 0$ is left *unclassified,* since the model either does not provide enough evidence, or provides conflicting evidence; fortunately it has been seen in all the real-life problems considered that the number of unclassified observations is extremely small.

4.3 Logical rating score (LRS)

In order to "learn" the Standard & Poor rating system, we shall proceed in two steps. In the first step, the proposed model is derived only from the information indicating for each pair of countries having different S&P ratings which of the two is rated higher. From this information, we shall derive an LAD model, whose discriminant provides a numerical measure $\Delta(P_{ij})$ of the "superiority" of the country i's rating over that of country j.

In the second step, applying multiple linear regression we derive new numerical ratings of all countries, called "logical rating scores" (LRS); the logical rating scores are obtained in such a way that their pairwise differences provide the best approximation of the numerical measures Δ obtained in the first step. These scores will be shown in Section 5 to have a very high correlation with the S&P ratings.

4.3.1 From pseudo-observations to relative preferences

The "observations" of the dataset used in the first step are those pseudo-observations P_{ij}, which correspond to countries i and j having different ratings. Each pseudo-observation P_{ij} is classified as positive or negative, according to the value of the indicator variable, i.e., depending on whether i is rated higher than j or vice versa. Clearly, the training set is anti-symmetric.

After having constructed the LAD model, we compute (according to (1.9)) the discriminant $\Delta(P_{ij})$ for each pseudo-observation P_{ij} ($i \neq j$). The values $\Delta(P_{ij})$ of the discriminant are called the *relative preferences,* and the [69 x 69]-dimensional anti-symmetric matrix, Δ, having them as components will be called the *relative preference matrix*. While the LAD model was derived using only those pseudo-observation P_{ij} for which i and j were rated differently, the discriminant matrix components are the values $\Delta(P_{ij})$ ($i \neq j$) taken by the discriminant for every pair of countries, including those that have the same S&P ratings.

To illustrate the concept of the relative preference matrix, let us consider the example of three countries, Japan, Canada, and Belgium, and the six associated pseudo-observations. The derived relative preferences obtained from the LAD model are shown in Table 9.

Table 9: Relative preferences for Japan, Canada and Belgium

	Japan	**Canada**	**Belgium**
Japan		0.00625	-0.00625
Canada	-0.00625		0.03125
Belgium	0.00625	-0.03125	

A naïve approach to the use of the relative preferences for deriving country ratings would rely on the direct interpretation of their sign as an indicator of rating superiority. More specifically, a large positive value $\Delta(P_{i,j})$ could be interpreted as country i being more creditworthy than country j, while the opposite conclusion could be drawn from a large negative value of the relative preference $\Delta(P_{i,j})$. A value equal or nearly equal to 0 would mean that the evidence for drawing conclusion about the relative creditworthiness of countries i and j is either lacking or conflicting.

The difficulty of this naïve approach is that it overlooks the imprecision of (i.e., the noise inherent in) data, and therefore of the relative preferences. The matrix above illustrates this phenomenon, since the naïve interpretation would rate Japan above Canada, Canada above Belgium, and at the same type Belgium above Japan, which contradicts the basic requirement of transitivity of an order relation.

In the following section in order to overcome this difficulty, we shall relax the overly constrained search for (possibly non-existent) country ratings whose pairwise orderings are in *precise* agreement with the signs of relative preferences, to the more flexible search for *logical rating scores* (*LRS*), having numerical values whose pairwise differences *approximate* well the relative preferences.

4.3.2 From relative preferences to logical rating scores using regression analysis

It has been common practice in the research literature (see e.g., Ferri and Liu, 1999, Hu et al., 2002, Kräussl, 2000, Monfort and Mulder, 2000, Sy, 2003) to interpret sovereign ratings as cardinal values. Assuming that the sovereign ratings β can be interpreted as cardinal values, it is natural to view the relative preferences Δ as differences of the corresponding ratings:

$$\Delta(P_{ij}) = \beta_i - \beta_j, \text{for all } i, j \in I, i \neq j \tag{1.10}$$

Obviously, the system (1.6) may or may not be consistent. We shall therefore replace it by:

$$\Delta(P_{ij}) = \beta_i - \beta_j + \varepsilon_{ij}, \text{for all } i, j \in I, i \neq j \tag{1.10}$$

The determination of those values of the β's which provide the best L_2 approximation of the Δ's can be found as a solution of the following multiple linear regression problem:

$$\Delta(\pi) = \sum_{k \in I} \beta_k * x_k(\pi) + \varepsilon(\pi) \quad , \tag{1.11}$$

where

$$\pi = \{(i, j) \mid i, j \in I, i \neq j\} \tag{1.12}$$

and

$$x_k(i,j) = \begin{cases} 1, & \text{for } k = i \\ -1, & \text{for } k = j \\ 0, & \text{otherwise} \end{cases} \tag{1.13}$$

The logical rating scores β_k estimated by the regression model are presented in column 4 of *Table 21* (in Appendix).

4.4 Evaluation of the results

The estimation of the regression model (1.6) used in deriving LRS shows that its statistical significance is very high. Indeed, its *p-value* turns out to be 0 and its *R*-square is 95.2%.

In order to evaluate the results obtained using the LRS model, we shall carry out several comparisons. We shall compare LRS with the S&P ratings, as well as with the results obtained using the non-recursive regression model derived from the S&P ratings. We shall further compare LRS with the scores associated to Moody's ratings at the end of December 1998, and with those provided by The Institutional Investor in March 1999. In this analysis, we shall evaluate first the relative preferences obtained using LAD, and then evaluate the LRS derived from these relative preferences. The consistency of any two sets of results is measured by their correlation coefficient.

4.4.1 Evaluation of relative preferences

In order to evaluate the matrix Δ of relative preferences obtained using LAD, we shall need a comparable point of reference. A natural benchmark of this sort can be associated to any set of numerical scores s_i representing sovereign ratings, by defining the *canonical relative preferences* d_{ij} to be simply the differences $d_{ij} = s_i - s_j$ associated to every pair of countries *i* and *j*. In this paper, we shall compare the LAD relative preferences Δ_{ij} with the canonical relative preferences $d^{S\&P}_{ij}$, d^{M}_{ij}, d^{II}_{ij}, d^{REG}_{ij}, and d^{LRS}_{ij} obtained respectively from the scores associated with S&P's ratings, Moody's ratings, The Institutional Investor's scores, the non-recursive regression model scores, and the logical rating scores. The corresponding matrices of relative preferences will be denoted $d^{S\&P}$, d^{M}, d^{II}, d^{REG}, and d^{LRS} respectively. The correlation levels are shown in Table 10, where the symmetric correlation matrix is filled out completely, in order to make comparisons easier.

Table: Correlation levels between relative preference matrices

	$d^{S\&P}$	d^{M}	d^{II}	d^{REG}	Δ	D^{LRS}
$d^{S\&P}$	100%	98.01%	96.18%	95.73%	93.21%	95.54%
d^{M}	98.01%	100%	96.31%	95.00%	92.89%	95.20%
d^{II}	96.18%	96.31%	100%	94.16%	91.82%	94.11%
d^{REG}	95.73%	95.00%	94.16%	100%	95.91%	98.29%
Δ	93.21%	92.89%	91.82%	95.91%	100%	97.57%
d^{LRS}	95.54%	95.20%	94.11%	98.29%	97.57%	100%

The high levels of correlation show that the relative preferences obtained using LAD are in a surprisingly good agreement both with the ratings of S&P and those of the other agencies, as well as with the non-recursive regression scores. A comparison of the logical rating scores with the LAD relative preferences shows the clear superiority of the former, in view of their higher degree of agreement with the other canonical relative preferences.

4.4.2 Evaluation of logical rating scores

We shall analyze now the correlation levels between the logical rating scores and the scores associated with the ratings of S&P, Moody and The Institutional Investor, as well as those provided by the non-recursive regression model (columns 2, 10 and 11, respectively, in *Table 21* in the Appendix). It can be shown (Appendix) that these correlation levels are exactly identical to those between the corresponding canonical relative preference matrices, which are presented in Table 10.

Several valuable conclusions can be derived from the correlation levels reported in Table 10. First, the high levels of correlation between the scores show that both the logical rating scores and those provided by the non-recursive regression model are very good approximations of the S&P ratings, as well as of those provided by other rating agencies.

Second, it is remarkable that, in spite of the entirely different nature of the LRS and the non-recursive regression techniques, their results have an extremely high correlation (98.29%). Moreover, the correlation between one of these two scores and any individual agency rating, is almost identical to the correlation between the other score and the rating of that agency.

This similarity is striking due to several essential distinctions between LRS and the non-recursive regression technique. The first distinction is that the regression technique assumes that the differences between consecutive ratings are the same, while the LRS is not based on any assumption about the magnitude of the differences. The second – and perhaps most striking distinction – concerns the mathematical techniques used for deriving the two scores: the non-recursive regression score is obtained through a classical statistical technique, while the LRS is derived through combinatorial-logical concepts and techniques.

The third distinction concerns the mathematical functions describing the relationship between the independent variables and the scores: while the non-recursive regression scores depend linearly on the variables, the LRS depends on them in a complex nonlinear fashion, through the intermediary constructs of logical patterns.

The most important conclusion derived from this discussion is the fact that the almost identical results, provided by two techniques of essentially different natures, strongly reinforce each other.

4.4.3 Discrepancies between S&P and LRS

It has been seen in the previous section that the LRS and the S&P ratings are in close agreement. However, since the logical rating scores and the S&P ratings are not expressed on the same scale, the comparison of the two scores of an individual country presents a challenge. In order to overcome this difficulty, we shall apply a linear transformation to the LRS, which brings them to the same scale as the S&P ratings. This is accomplished by determining the coefficients a and c for the transformation $a*\beta_i + c$ of the LRS β_i in such a way that the mean square difference between the transformed LRS and the S&P ratings is minimized. Obviously, these coefficients can be determined by simple linear regression. As a result of this transformation, the LRS become directly comparable with the S&P ratings. Clearly, the consistency of (i.e., the correlation coefficient between) the LRS and S&P ratings is not affected by this transformation.

Using formula (1.6), we then compute the confidence intervals for the transformed LRS of each country in the dataset. In 1998, five countries have a S&P rating that does nor fall within the confidence interval of the transformed LRS. Columbia appears to be too favorably rated by S&P, while Hong-Kong, Malaysia, Pakistan and Russia appear to be rated too harshly by S&P. As explained in Sections 3.1.3 and 3.1.4, the evolution of the S&P ratings for Columbia, Pakistan and Russia is in agreement with the 1998 LRS of these countries, and underlines the prediction capability of the LRS model. It is remarkable that the evolution of the S&P ratings of Malaysia and Hong-Kong is also in agreement with their 1998 LRS. Indeed, both Malaysia and Hong-Kong have been upgraded shortly thereafter, the former moving from BBB- to BBB in November 1999, and the latter from A to A+ in February 2001.

4.5 Extendability of the logical rating scores model

As it has been seen above, the derivation of the LRS model consists of two stages. In the first stage, LAD is applied to derive patterns of superiority of one country over another and combine them into a discriminant that provides numerical value of the relative preferences. In the second stage, the relative preferences are "integrated" into logical rating scores, by using classical linear regression.

The LRS model makes possible the calculation of logical rating scores on the basis of new data. Indeed, in order to obtain logical rating scores for the new data all that has to be done is to recalculate the relative preferences using the existing LAD discriminant, and then to rerun linear regression in order to "reintegrate" the new relative preferences into the new logical rating scores. Clearly, the analysis of new data by the LRS model is carried out independently of any possible new ratings of S&P, since the LAD discriminant underlying the LRS model has already been constructed and remains valid.

The remarkable feature of the LRS model is that for important cases of data perturbation, including those resulting from temporal changes in the values of the independent variables, or the addition of previously non-rated countries, the new logical rating scores maintain to a large extent their validity. In order to illustrate this idea, we shall show below that the LRS model developed above on the basis of the 1998 data can be successfully applied on the one hand for deriving LRS scores based on 1999 data, and on the other hand for deriving LRS scores of previously non-rated countries.

4.5.1 Temporal changes in data

In this section, we shall construct logical rating scores based on the 1999 data, in order to evaluate the robustness and consistency of the LRS model. This goal will be accomplished by comparing the LRS obtained in this way with Standard and Poor's 1999 ratings. The logical ratings scores for the 1999 data are given in Appendix (*Table 21*, columns 8).

It has been seen in Section 4.4.2 that the scores provided by the two models (the non-recursive regression and the LRS models) built on the 1998 Standard & Poor ratings when applied to the 1998 data are strikingly similar. In this section, we compare the results of the same two methods applied to the 1999 data.

Table 11: Correlation levels between relative preference matrices

	$d^{S\&P}$	d^{REG}	Δ	d^{LRS}
$d^{S\&P}$	100%	94.74%	91.61%	94.12%
d^{REG}	94.74%	100%	94.46%	97.50%
Δ	91.61%	94.46%	100%	96.48%
d^{LRS}	94.12%	97.50%	96.48%	100%

The first conclusion resulting from the high levels of pairwise correlations between the S&P 1999 ratings, the relative preferences given by the LAD discriminant, and the canonical relative preferences corresponding to non-recursive regression scores and LRS, is that both the LRS and the non-recursive regression model have a very strong temporal stability. The second conclusion is that the logical rating scores are superior to the relative preferences given by the LAD discriminant.

The major conclusion resulting from this table is that the logical rating and the non-recursive regression scores remain strikingly similar, and each of them provides a strong reinforcement of the other.

4.5.2 Discrepancies

Using formula (1.6), we compute the confidence intervals for the transformed LRSe of each country in the dataset. Applying the 1998 LRS model to the 1999 data, we see that only two countries (Russia and Hong-Kong) have S&P ratings that are outside the confidence intervals of the corresponding transformed LRS. These two countries appear to be rated too harshly by S&P. As explained in Sections 3.1.4 and 4.4.3, the evolution of the S&P ratings for Hong-Kong and Russia is in agreement with the 1999 LRS of these countries.

4.5.3 Previously non-rated countries

The availability of the LAD discriminant, which does not involve in any way the previous years' S&P ratings, makes it possible to rate previously non-rated countries in the following way. After calculating first the attribute values of all the pseudo-observations involving the new countries to be evaluated, the relative preferences are to be calculated for these pseudo-observations, and the resulting columns and rows have to be added to the matrix of relative preferences. The new LRS for all the countries (new and old) should then be determined by running the multiple linear regression model (1.12).

In order to evaluate the capability of LRS to correctly predict S&P ratings, we compare the LRS predicted as described above, with the S&P ratings when they first become available. The direct comparison between these ratings is carried out using the linear transformation described in Section 4.4.3.

Using formula (1.6), we compute the confidence intervals for the transformed LRS of four countries never rated by S&P by December 1998. It appears that our predictions for three of them (Guatemala, Jamaica and Papua New Guinea) correspond perfectly to the first time (subsequent) S&P ratings. The comparison between the LRS and the first S&P rating (SD) given in July 2000 for the fourth country (Ecuador) shows that S&P rated it much too harshly, since one month later S&P significantly raised its rating to B-, thus fully justifying the LRS prediction.

4.6 Cross-validation of relative preferences

Since the role of the LAD discriminant is to capture the rules of country creditworthiness, which are implicit in the S&P ratings, it is the most important component of the LRS methodology. As any learning procedure, LAD can be susceptible to overfitting, i.e., adapting so well to the training data to allow random noise to influence the model. If this happens, the performance of the resulting model can be excellent on the training data, but can perform very poorly on new observations.

To test whether the LAD discriminant is affected by overfitting, we shall use – similarly to Section 3.1.2– the commonly used statistical technique of cross-validation. The particular type of cross-validation technique used here is known as "jackknife" (Quenouille, 1949,1956) or "leave-one-out". In broad terms, the jackknife technique consists of removing from the dataset one observation at a time, learning a model from all the remaining observations, evaluating the resulting model on the removed observation; and then repeating these steps for each observation in the dataset. If the predicted evaluations are "close to" the actual values of the observations, as well as in-the-sample evaluations, then the model is not affected by overfitting.

Each step of the implementation of jackknife used here eliminates from the dataset all the pseudo-observations involving a certain country, derives the LAD discriminant on the basis of the remaining pseudo-observations, and then uses this discriminant to evaluate the relative preference for every removed pseudo-observation. This provides a column of relative preferences of pseudo-observations involving the country singled out at this step. After repeating this procedure for every country in the dataset, we combine the obtained columns into a matrix of relative preferences denoted by Δ^{JK}. Then we construct the logical rating scores on the basis of Δ^{JK} and denote the corresponding matrix of canonical relative preferences by d_{JK}^{LRS}.

Table12: Correlation between cross-validated relative preferences

	$d^{S\&P}$	Δ	d^{LRS}	Δ^{JK}	d_{JK}^{LRS}
$d^{S\&P}$	100%	93.21%	95.54%	92.98%	95.26%
Δ	93.21%	100%	97.57%	96.48%	96.36%
d^{LRS}	95.54%	97.57%	100%	96.35%	96.89%
Δ^{JK}	92.98%	96.48%	96.35%	100%	97.43%
d_{JK}^{LRS}	95.26%	96.36%	96.89%	97.43%	100%

The correlation levels presented in Table 12 indicate that the matrices of relative preferences Δ^{JK} and d_{JK}^{LRS} obtained through the jackknife procedure are highly correlated with (i) the canonical relative preferences corresponding to the S&P ratings $d^{S\&P}$, and (ii) the in-the-sample relative preferences Δ, as well as the corresponding logical rating scores d^{LRS} . This shows that Δ and d^{LRS} are not affected by overfitting.

5 Concluding Remarks

The central objective of this paper was to develop transparent, consistent, self-contained, and stable country risk rating systems, closely approximating the country risk ratings provided by a major rating agency (Standard & Poor). We proposed in this paper two models that achieve the stated objectives, the first one utilizing the classical econometric technique of multiple linear regression, and the second one using the combinatorial-logical technique of Logical Analysis of Data.

The proposed models are highly *accurate*, having a 95.5% correlation level with the actual S&P ratings and almost equally high correlations with Moody's and The Institutional Investor. The models avoid overfitting, as demonstrated by the 99.1% correlation between in- and out-of-sample rating predictions, calculated by *k*-fold cross-validation. The proposed models are *transparent* since they make the role of the economic-financial and political variables explicit.

The proposed models are distinguished from the rating models in the existing literature by their *self-contained* nature, i.e., by their non-reliance on any information derived from lagged ratings. Therefore, the high level of correlation between predicted and actual ratings cannot be attributed to the reliance on lagged ratings, and is a reflection of the relevance and predictive power of the independent variables included in these models. The significant advantage of the non-recursive nature of the proposed models is their applicability to not-yet-rated countries.

The *consistency* of the proposed models is illustrated by the fact that the few discrepancies between the S&P ratings and those of the proposed models were resolved by subsequent changes in S&P's ratings. The *stability* of the constructed non-recursive models is shown in two ways: by predicting 1999 ratings using the non-recursive models derived from the 1998 dataset applied to the 1999 data, and – most importantly -- by successfully predicting the ratings of several previously non-rated countries, i.e., in full agreement with subsequent ratings of those countries by S&P.

The confidence in the results and in the validity of both models is strongly reinforced by the fact that the traditional linear regression model and the qualitatively different combinatorial-logical model produce almost identical results.

The significance of the results of this paper is further confirmed in a forthcoming study (Hammer et al., 2004), in which the LAD discriminant constructed in this paper is used for deriving a partial order representing the creditworthiness of countries, and then extending this partial order to a new rating system. The comparison of the results of the present and the forthcoming studies shows that – in spite of yet another qualitatively different data analysis approach -- the predicted ratings are strongly correlated, at the surprising level of 98-99%.

References

Altman E.I. and Saunders A. 1997. Credit Risk Measurement: Developments over the Last 20 Years. *Journal of Banking and Finance* **21** (11-12), 1721-1742.

Aylward L. and Thorne R. 1998. Countries' Repayments Performance Vis-à-Vis the IMF", *IMF Staff Papers* **45** (4), 595-619.

Bhatia A.V. 2002. Sovereign Credit Risk Ratings: An Evaluation. *IMF Working Pape* ***WP/*** **03/170**.

Boros E., Hammer P.L., Ibaraki T. and Kogan A. 1997. Logical Analysis of Numerical Data, *Mathematical Programming* **79**, 163-190.

Bouchet M.H., Clark E. and Groslambert B. 2003. Country Risk Assessment: A Guide to Global Investment Strategy. John Wiley & Sons Ltd. Chichester, England.

Bourke P. and Shanmugam B. 1990. *An Introduction to Bank Lending.* Addison-Wesley Business Series.

Breiman L. and Spector P. 1992. Submodel Selection and Evaluation in Regression: The X-Random Case. *International Statistical Review* **60**, 291-319.

Brewer T.L. and Rivoli P. 1990. Politics and Perceived Country Creditworthiness in International Banking. *Journal of Money, Credit and Banking* **22**, 357-369.

Brewer T.L. and Rivoli P. 1997. Political Instability and Country Risk. *Global Finance Journal* **8** (2), 309-321.

Cantor R. and Packer F. 1996. Determinants and Impact of Sovereign Credit Ratings. *FRBNY Economic Policy Review,* 37-53.

Citron J.T. and Neckelburg G. 1987. Country Risk and Political Instability. *Journal of Development Economics* **25**, 385-395.

Claessens S. and Embrechts G. 2002. Basel II, Sovereign Ratings and Transfer Risk: External versus Internal Ratings. Presentation at the *Basel II: An Economic Assessment, Bank for International Settlements*, Basel, Switzerland.

Cook W.D. and Hebner K.J. 1993. A Multicriteria Approach to Country Risk Evaluation: With an Example Employing Japanese Data. *International Review of Economics and Finance* **2** (4), 327-348.

Coplin W.D. and O'Leary M.K. 19983. *Political Risk Yearbook*, The PRS Group, New York.

Cosset J.C. and Roy J. 1991. The Determinants of Country Risk Ratings. *Journal of International Business Studies* **22**, 135-142.

Cosset J.C., Siskos Y. and Zopounidis C. 1992. Evaluating Country Risk: A Decision Support Approach. *Global Finance Journal* **3**, 79-95.

Crama Y., Hammer P.L. and Ibaraki T. 1988. Cause-Effect Relationships and Partially Defined Boolean Functions. *Annals of Operations Research* **16**, 299-326.

Cruces J.J. 2003. Statistical Properties of Sovereign Credit ratings. *Working Paper*. University of San Andres, Buenos Aires.

Doumpos M. and Zopounidis C. 2001. Assessing Financial Risks Using a Multicriteria Sorting Procedure: The Case of Country Risk Assessment. *Omega* **29**, 97-109.

Easton S.T. and Rockerbie D.W. 1999. What's in a Default ? Lending to LDCs in the Face of Default Risk. *Journal of Development Economics* **58** (2), 319-332.

Eaton J., Gersovitz M. and Stiglitz J.E. 1986, The Pure Theory of Country Risk. *European Economic Review* **30**, 481-513.

Efron B. 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, SIAM.

Erb C.B., Harvey C.R. and Viskanta T.E. 1996a. Expected Returns and Volatility in 135 Countries. *Journal of Portfolio Management*, 46-58.

Erb C.B., Harvey C.R. and Viskanta T.E. 1996b. Political Risk, Economic Risk and Financial Risk. *Financial Analysts Journal*, 29-46.

Estrella A. 2000. Credit Ratings and Complement Sources of Credit Quality Information. *Basel Committee on Banking Supervision Working Paper* **3.**

Feder G. and Uy V.U. 1985. The Determinants of International Creditworthiness and their Policy Implications. *Policy Modeling* **7** (1), 133-156.

Ferri G., Liu L-G. and Stiglitz J. 1999. The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis. *Economic Notes* **3**, 335-355.

Hammer P.L. 1986. Partially Defined Boolean Functions and Cause-Effect Relationships. *International Conference on Multi-Attribute Decision Making Via OR-Based Expert Sytems*. University of Passau, Passau, Germany.

Hammer P.L., Kogan A. and Lejeune M.A. 2004. Modeling Country Risk Ratings Using Partial Orders. *RUTCOR Technical Report* **1-2004**, Rutgers University, New Brunswick, New Jersey.

Haque N.U., Kumar M.S., Mark N. and Mathieson D. 1996. The Economic Content of Indicators of Developing Country Creditworthiness. *International Monetary Fund Working Paper* **43** (4), 688-724.

Haque N.U., Kumar M.S., Mark N. and Mathieson D. 1998. The Relative Importance of Political and Economic Variables in Creditworthiness Ratings. *International Monetary Fund Working Paper* **46**, 1-13.

Haque N.U., Mark N. and Mathieson D. 1997. Rating the Raters of Country Creditworthiness. *Finance & Development* **34**, 10-13.

Hjorth J.S.U. 1994. *Computer Intensive Statistical Methods Validation, Model Selection, and Bootstrap*. London, Chapman & Hall.

Hu Y.-T., Kiesel R. and Perraudin W. 2002. The Estimation of Transition Matrices for Sovereign Credit Ratings. *Journal of Banking and Finance* **26** (7), 1383-1406.

Hurvich C.M. and Tsai C.-L. 1989. Regression and Time Series Model Selection in Small Samples. *Biometrika* **76**, 297-307.

Jüttner J.D. and McCarthy J. 2000. Modeling a Rating Crisis. Macquarie University, Unpublished, Sidney, Australia, http://www.econ.mq.edu.au/staff/djjuttner/SOVEIG2.pdf.

Kaminsky G. and Schmukler S.L. 2002. Emerging Market Instability: Do Sovereign Ratings Affect Country Risk and Stock Returns? *World Bank Economic Review* **16**, 171-195.

Kaufmann D., Kraay A. and Zoido-Lobaton P. 1999a. Aggregating Governance Indicators. *World Bank Policy Research Department Working Paper* **2195**, http://www.worldbank.org/wbi/governance/pdf/agg_ind.pdf.

Kaufmann D., Kraay A. and Zoido-Lobaton P. 1999b. Governance Matters. *World Bank*

Policy Research Department Working Paper **2196**, http://www.worldbank.org/wbi/governance/pdf/govmatrs.pdf.

Kaufmann D., Kraay A. and Zoido-Lobaton P. 2002. Governance Matters II: Updated Indicators for 2000/01. *World Bank Policy Research Department Working Paper* **2772**, http://ideas.repec.org/p/wop/wobago/2772.html.

Kräussl R. 2000. Sovereign Ratings and Their Impact on Recent Financial Crises. *Working paper, Center For Financial Studies*. Frankfurt.

Kunczik M. 2000. Globalization: News Media, Images of Nations and the Flow of International Capital with Special Reference to the Role of Rating Agencies, Paper presented at the IAMCR Conference, Singapore, 1-49.

Larrain G., Reisen H. and von Maltzan J. 1997. Emerging Market Risk and Sovereign Credit Ratings. *OECD Development Center*, 1-30.

Lee S.H. 1993. Relative Importance of Political Instability and Economic Variables on Perceived Country Creditworthiness. *Journal of International Business Studies* **24**, 801-812.

Levich R., Reinhart C., and Majnoni G. 2002. *Ratings, Rating Agencies and the Global financial System.* Kluwer Academic Press. New York.

Manasse P., Roubini N. and Schimmelpfennig A. 2003. Predicting Sovereign Debt Crises. *IMF Working Pape* ***WP/*** **03/221**.

Mathieson D. and Schinasi G. 1999. International Capital Markets: Developments, Prospects, and Key Policy Issues. *IMF Series: World Economic and Financial Surveys*.

Mauro P. 1993. Essays on Country Risk, Asset Markets and Economic Growth. Harvard University Dissertation.

Mauro P. 1998. Corruption and the Composition of Government Expenditure. *Journal of Public Economics* **69** (2), 263-279.

Monfort B. and Mulder C. 2000. Using Credit Ratings for Capital Requirements on Lending to Emerging Market Economies: Possible Impact of a New Basel Accord. *IMF Working Paper* WP/00/69.

Moody's. 2001. *Moody's Country Credit Statistical Handbook*, First Edition, New York..

Mulder C. and Perelli R. 2001. Foreign Currency Credit Ratings for Emerging Market Economies. *IMF Working Paper* **WP/01/191**.

Mulekar M.S. and Mishra S.N. 2000, Confidence Interval Estimation of Overlap: Equal Means Case, *Computational Statistics & Data Analysis* **34**, 121-137.

Bilson C. M., Brailsford T.J. and Hooper V.J. 2001. Selecting Macroeconomic Variables as Explanatory Factors of Emerging Stock Market Returns. *Pacific-Basin Finance Journal* **9** (4), 401-426 .

Neter J., Kutner M.H., Nachtsheim C.J. and Wasserman W. 1996. Applied Linear Statistical Models. Fourth Edition. Irwin/McGraw-Hill.

Nickell P., Perraudin W. and Varotto S. 2000. Stability of Rating Transitions. *Journal of Banking and Finance* **24**, 203-227.

Nordal K.B. 2001. Country Risk, Country Risk Indices and Valuation of FDI: A Real Options Approach. *Emerging Markets Review* **2**, 197-217.

Oral M., Kettani O., Cosset J.-C. and Daouas M. 1992. An Estimation Model for Country Risk Rating. *International Journal of Forecasting* **8** (4), 583-593.

Quenouille, M. 1949. Approximate Tests of Correlation in Time Series. *Journal of the Royal Statistical Society, Series B* **11**, 18-84.

Quenouille, M. 1956. Notes on Bias in Estimation. *Biometrika* **43**, 353-360.

Reinhart C.M. 2002. Default, Currency Crises, and Sovereign Credit Ratings. *World Bank Economic Review* **16**, 151-170.

Reisen H. and von Maltzan J. 1998. Sovereign Credit Ratings, Emerging Market Risk and Financial Market Volatility. *Intereconomics* **33** (2), 73-82.

Reisen H. and Von Maltzan J. 1999. Boom and Bust and Sovereign Ratings. *International Finance* **2** (2), 273-293.

Saini K.G. and Bates P.S. 1984. A Survey of the Quantitative Approaches to Country Risk Analysis. *Journal of Banking and Finance* **8**, 341-356.

Shao J. 1993. Linear Model Selection by Cross-Validation. *Journal of the American Statistical Association* **88**, 486-494.

Shao J. and Tu D. 1995. *The Jackknife and Bootstrap*. New York, Springer-Verlag.

Simpson J. 2000. The World Trade Organisation and Basel Accord Membership and Compliance in Relation to International Banking Risk. *Curtin University Working Paper* **35.**

Somerville R.A. and Taffler R.J. 1995. Banker Judgment versus Formal Forecasting Models : The Case of Country Risk Assessment. *Journal of Banking and Finance* **19** (2), 281-297.

Standard & Poor. 1999. Sovereign Ratings Display Stability over Two Decades. *Credit Week.*

Standard & Poor. 2000. Ratings Performance. Standard & Poor. New York.

Standard & Poor. 2001. *Sovereign Ratings History since 1975.* New York.

Sy A.N.R. 2003. Rating the Rating Agencies: Anticipating Currency Crises or Debt Crises? *IMF Working Paper* ***WP/*** **03/122**.

Tang J.C.S. and Espinal C.G. 1989. A Model to Assess Country Risk. *OMEGA: International Journal of Management Sciences* **17** (4), 363-367.

World Bank. 2000a. *Global Development Finance 2000.* Washington D.C.

World Bank. 2000b. *World Development Indicators 2000.* Washington D.C.

Appendix

Table 13:Economic variables and literature

Variable	Literature
Consumer price index	Larrain et al. (1997), Hu et al. (2002)
Credit claims on central government growth rate	Monfort and Mulder (2000)
Current account balance / GDP	Haque et al. (1996,1998), Larrain et al. (1997), Brewer and Rivoli (1990), Doumpos et al. (2001), Cosset et al.(1992), Cook and Hebner (1993), Tang and Espinal (1990)
Debt[12] / exports	Larrain et al. (1997), Feder and Uy (1985), Aylward and Thorne (1998), Dailami and Leipziger (1997), Cosset and Roy (1991), Cosset et al. (1992), Cantor and Packer (1996), Lee (1993), Doumpos and Zopounidis (2001), Monfort and Mulder (2000), Hu et al. (2002), Mulder and Perelli (2002)
Debt[13] / GDP	Haque et al. (1996, 1998), Feder and Uy (1985), Lee (1993), Brewer and Rivoli (1990), Aylward and Thorne (1998), Doumpos and Zopounidis (2001), Cook and Hebner (1993), Hu et al. (2002)
Debt / reserves	Monfort and Mulder (2000), Manasse et al. (2003)
Dependence on oil exportation	Feder and Uy (1985)
Domestic investment / GDP	Larrain et al. (1997), Monfort and Mulder (2000)
Exports / GDP	Aylward and Thorne (1998)
Exports concentration	Feder and Uy (1985)
Exports growth rate	Haque et al. (1996,1998), Feder and Uy (1985), Doumpos and Zopounidis (2001), Cosset et al.(1992), Monfort and Mulder (2000)
Exports variability	Cosset et al.(1992)
Exports vulnerability to external shocks	Feder and Uy (1985)
External debt / GDP	Larrain et al. (1997}, Brewer and Rivoli (1990), Monfort and Mulder (2000), Manasse et al. (2003)
Fiscal balance[14]	Larrain et al. (1997), Cantor and Packer (1996), Lee (1993), Monfort and Mulder (2000) , Cook and Hebner (1993)
Foreign investment policy	Cook and Hebner (1993)
GDP[15] growth rate	Haque et al. (1996,1998), Larrain et al. (1997), Feder and Uy (1985), Cantor and Packer (1996), Doumpos and Zopounidis (2001), Monfort and Mulder (2000), Cook and Hebner (1993), Hu et al.(2002)
GDP per capita	Dailami and Leipziger (1997), Erb et al. (1997), Feder and Uy (1985), Cosset et al. (1991, 1992), Larrain et al. (1997), Monfort and Mulder (2000), Tang and Espinal (1990)
GDP per capita growth rate	Lee (1993), Haque et al. (1996), Aylward and Thorne (1998)
Gross investment / GDP	Easton and Rockerbie (1999), Cosset et al. (1991, 1992), Doumpos and Zopounidis (2001)

[12] Same as for footnote 6.
[13] The word "debt" can encompass foreign, total, debt service or external debt, depending on authors.
[14] Central government spending / GDP, domestic public debt / GDP and are used as a proxy for this variable.
[15] Note that authors use GDP as well as GNP.

Imports / GDP	Haque et al.(1996), Aylward and Thorne (1998)
Imports growth rate	Doumpos and Zopounidis (2001)
Income velocity of money (GDP/M2)	Doumpos and Zopounidis (2001)
Indicator for economic development	Cantor and Packer (1996)
Inflation rate	Haque et al. (1996,1998), Larrain et al. (1997), Dailami and Leipziger (1997), Erb et al. (1997), Aylward and Thorne (1998), Cantor and Packer (1996), Doumpos and Zopounidis (2001), Monfort and Mulder (2000)
International reserves / imports	Haque et al. (1996,1998), Easton and Rockerbie et al. (1999), Dailami and Leipziger (1997), Feder and Uy (1985), Aylward and Thorne (1998), Lee (1993), Doumpos and Zopounidis (2001), Cosset et al.(1992), Monfort and Mulder (2000), Tang and Espinal (1990), Hu et al.(2002)
Long-term debt / GDP	Easton and Rockerbie et al. (1999)
Oil exportation[16]	Feder and Uy (1985)
Real exchange rate	Haque et al. (1996,1998), Larrain et al.(1997), Monfort and Mulder (2000), Cook and Hebner (1993)
Savings / GDP	Larrain et al. (1997)
Short-term debt / reserves	Dailami and Leipziger (1997)
Short-term debt / total debt	Monfort and Mulder (2000)
Terms of trade	Haque et al. (1996,1998), Easton and Rockerbie (1999), Feder and Uy (1985), Doumpos and Zopounidis (2001), Monfort and Mulder (2000)
Trade openness	Easton and Rockerbie (1999)
Treasury bill rate	Haque et al. (1998), Monfort and Mulder (2000)

Table 14: Political variables and literature

Variable	**Literature**
Anti-governmental demonstrations	Haque et al. (1998)
Armed conflicts (or riots)	Haque et al. (1998), Brewer and Rivoli (1990), Cook and Hebner (1993)
Assassination	Haque et al. (1998)
Corruption	Mauro (1993)
Coups	Haque et al. (1998)
General strikes	Haque et al. (1998)
Guerilla warfare	Haque et al. (1998)
Influence of the middle class	Mauro (1993), Cook and Hebner (1993)
Legal system	Mauro (1993)
Major government crises	Haque et al. (1998)
Political change	Mauro (1993), Brewer and Rivoli (1990)
Political legitimacy	Brewer and Rivoli (1990)
Political stability	Brewer and Rivoli (1990, 1997), Feder and Uy (1985), Citron and Neckelburg (1987), Mauro (1993), Lee (1993), Cosset et al.(1992),

[16] Represented by a dummy variable.

	Cook and Hebner (1993), Manasse et al. (2003)
Probability of opposition group takeover	Mauro (1993)
Purges	Haque et al. (1998)
Red tape, bureaucracy	Mauro (1993)
Relationships with neighboring countries	Mauro (1993)
Revolutions	Haque et al. (1998)
Social Stability	Cook and Hebner (1993)
Stability of labor	Mauro (1993)
Terrorism	Mauro (1993)

Table 15: Standard & Poor's country rating system

	Level	**Description**
INVESTMENT RATING	**AAA**	An obligor rated AAA has extremely strong capacity to meet its financial commitments. AA is the highest issuer credit rating assigned by S&P.
	AA	An obligor rated AA has very strong capacity to meet its financial commitments. It differs from the highest rated obligors only in small degree.
	A	An obligor rated A has strong capacity to meet its financial commitments but is somewhat more susceptible to the adverse effects of changes in circumstances and economic conditions than obligors in higher-rated categories.
	BBB	An obligor rated BBB has adequate capacity to meet its financial commitments. However, adverse economic conditions or changing circumstances are more likely to lead to a weakened capacity of the obligor to meet its financial commitments.
SPECULATIVE RATING	**BB**	An obligor rated BB is less vulnerable in the near term than other lower-rated obligors. However, it faces major ongoing uncertainties and exposure to adverse business, financial, or economic conditions which could lead to its inadequate capacity to meet financial commitments.
	B	An obligor rated B is more vulnerable than the obligors rated BB, but, at the time of the rating, it has the capacity to meet financial commitments. Adverse business, financial, or economic conditions could likely impair its capacity or willingness to meet financial commitments.
DEFAULT RATING	**CCC**	An obligor rated CCC is vulnerable at the time of the rating, and is dependent upon favorable business, financial, and economic conditions to meet financial commitments.
	CC	An obligor rated CC is highly vulnerable at the time of the rating.
	C	An obligor rated C is vulnerable to nonpayment at the time of the rating and is dependent upon favorable business, financial, and economic conditions to meet financial commitments.
	D	An obligator rated D is predicted to default.
	SD	An obligator rated SD (selected default) is presumed to be unwilling to repay.

Table 16: Standard & Poor's country ratings (end of December, 1998)

We have converted the Standard & Poor rating scale (columns 1 and 4) into a numerical scale (columns 2 and 5). Such a conversion is not specific to us. Bouchet et al. (2003), Estrella (2000), Ferri et al.(2001), Kräussl (2000), Monfort and Mulder (2000), Mulder and Perelli (2001), Sy [2003] proceed similarly. Moreover, Bloomberg, a major provider of financial data services, developed a standard cardinal scale for comparing Moody's, S&P and Fitch-BCA ratings (Kaminsky and Schmukler, 2002). A higher numerical value denotes a higher probability of default. The numerical scale is referred to in this paper as Standard & Poor's preorder.

Rating	Preorder	Country	Rating	Preorder	Country
AAA	21	AUSTRIA	BBB-	12	CROATIA
AAA	21	FRANCE	BBB-	12	EGYPT
AAA	21	GERMANY	BBB-	12	LITHUANIA
AAA	21	JAPAN	BBB-	12	MALAYSIA
AAA	21	NETHERLANDS	BBB-	12	POLAND
AAA	21	NORWAY	BBB-	12	THAILAND
AAA	21	SINGAPORE	BBB-	12	TUNISIA
AAA	21	SWITZERLAND	BBB-	12	URUGUAY
AAA	21	UNITED KINGDOM	BB+	11	KOREA
AAA	21	UNITED STATES	BB+	11	PANAMA
AA+	20	BELGIUM	BB+	11	PHILIPPINES
AA+	20	CANADA	BB+	11	SLOVAK REPUBLIC
AA+	20	DENMARK	BB+	11	SOUTH AFRICA
AA+	20	IRELAND	BB+	11	TRINIDAD AND TOBAGO
AA+	20	NEW ZEALAND	BB	10	ARGENTINA
AA+	20	SWEDEN	BB	10	COSTA RICA
AA	19	AUSTRALIA	BB	10	EL SALVADOR
AA	19	FINLAND	BB	10	INDIA
AA	19	ITALY	BB	10	MEXICO
AA	19	PORTUGAL	BB	10	MOROCCO
AA	19	SPAIN	BB	10	PERU
A+	17	CYPRUS	BB-	9	BOLIVIA
A+	17	ICELAND	BB-	9	BRAZIL
A+	17	MALTA	BB-	9	JORDAN
A	16	HONG KONG	BB-	9	LEBANON
A	16	SLOVENIA	BB-	9	PARAGUAY
A-	15	CHILE	B+	8	DOMINICAN REPUBLIC
A-	15	CZECH REPUBLIC	B+	8	KAZAKHSTAN
A-	15	ISRAEL	B+	8	VENEZUELA
BBB+	14	CHINA	B	7	TURKEY
BBB+	14	ESTONIA	B-	6	ROMANIA
BBB	13	GREECE	CCC+	5	INDONESIA
BBB	13	HUNGARY	CCC-	3	RUSSIA
BBB	13	LATVIA	CC	2	PAKISTAN
BBB-	12	COLUMBIA			

Table 17: Ratings with the non-recursive regression model approach

Countries	Non-recursive regression model: predicted country risk ratings and 10-fold cross-validation: Predicted country risk rating (in-the-sample)	Cross-validation country risk rating (out-of-the-sample)	Country risk ratings based on economic variables	Predicting 1999 country risk ratings using 1998 model	Predicting 1999 country risk ratings using 1999 model
Argentina	12.09	14.01	13.25	14.19	13.32
Australia	19.92	20.85	20.23	21.37	18.19
Austria	20.62	19.51	20.13	20.52	18.43
Belgium	17.91	16.55	18.32	18.44	17.53
Bolivia	8.88	8.84	8.73	8.88	7.95
Brazil	9.62	10.17	10.37	9.77	7.96
Canada	20.20	20.18	20.53	20.85	18.90
Chile	14.99	14.63	14.80	15.15	12.96
China	13.37	13.06	11.44	13.37	12.47
Colombia	17.61	18.88	18.45	18.59	17.63
Costa Rica	8.43	7.38	8.84	8.74	8.11
Croatia	11.82	12.06	12.36	12.42	9.28
Cyprus	11.89	11.62	11.52	11.36	10.49
Czech Republic	17.16	17.11	17.46	18.14	16.68
Denmark	15.04	15.03	14.45	15.44	13.44
Dominican Rep	20.30	20.30	20.13	20.73	18.45
Egypt	8.82	8.78	8.69	9.65	8.47
El Salvador	10.73	10.38	10.01	10.71	10.37
Estonia	9.99	9.88	9.60	10.48	9.72
Finland	12.07	11.92	13.34	12.60	10.82
France	19.28	19.20	19.74	20.00	17.03
Germany	18.42	18.27	18.97	19.18	17.14
Greece	20.34	20.22	20.44	20.96	18.41
Hong-Kong	13.44	13.59	14.25	14.45	12.32
Hungary	13.28	13.01	13.29	13.48	9.33
Iceland	20.39	20.65	21.03	21.40	19.18
India	8.05	7.48	9.40	8.62	6.29
Indonesia	5.28	5.48	6.63	6.75	4.23
Ireland	18.91	19.11	20.09	20.49	18.35
Israel	14.00	14.19	14.53	14.20	14.60

Italy	17.30	17.52	16.90	17.36	15.78
Japan	20.07	19.85	19.62	20.08	17.58
Jordan	9.95	9.73	11.21	10.57	7.73
Kazakhstan	8.57	8.41	7.92	9.00	6.83
Korea. Rep.	14.36	14.14	14.54	15.39	12.27
Latvia	11.20	10.82	11.53	11.32	9.55
Lebanon	9.86	10.41	10.07	9.34	9.32
Lithuania	10.57	10.03	12.12	10.62	9.03
Malaysia	14.48	15.00	13.83	14.23	11.45
Malta	15.87	15.73	16.03	16.13	13.54
Mexico	10.35	10.15	9.97	10.95	8.40
Morocco	10.38	10.70	10.33	10.20	8.97
Netherlands	20.62	20.56	21.04	21.01	17.85
New Zealand	19.81	19.49	18.96	19.66	16.93
Norway	22.60	22.92	20.72	22.40	21.10
Pakistan	5.67	5.94	7.16	6.24	4.07
Panama	8.87	10.69	10.39	11.52	9.60
Paraguay	7.46	7.55	6.57	7.48	8.45
Peru	10.57	10.88	9.60	10.25	9.81
Philippines	10.58	10.94	9.75	9.94	6.36
Poland	12.95	12.84	12.68	12.90	9.93
Portugal	17.57	17.37	16.87	17.44	14.91
Romania	7.93	8.97	5.49	8.32	5.72
Russia	6.37	7.27	1.59	6.49	4.04
Singapore	20.12	19.58	19.23	18.54	17.08
Slovak Republic	11.03	11.30	11.85	12.74	10.52
Slovenia	14.19	13.80	14.80	15.28	12.87
South Africa	11.35	11.57	12.04	12.87	11.08
Spain	17.66	17.75	18.31	18.24	15.40
Sweden	19.26	19.15	19.42	19.54	17.52
Switzerland	23.69	24.07	22.81	23.76	21.46
Thailand	12.15	12.54	11.94	13.13	10.34
Trinidad & Tob	11.04	10.75	12.61	12.68	10.41
Tunisia	11.59	11.65	13.58	12.92	8.60
Turkey	5.07	2.94	4.81	6.41	4.52
UK	20.27	20.15	20.44	20.77	18.56
United States	21.08	23.05	22.07	23.98	22.90
Uruguay	13.31	13.25	13.63	14.12	10.43
Venezuela	7.44	6.79	6.55	8.42	6.07

Table 18: Moody's rating system

	Levels	Meaning		Levels	Meaning
INVESTMENT RATING	**Aaa**	Highest quality	**SPECULATIVE RATING**	**Ba1**	Likely to fulfill obligations
	Aa1	High quality		**Ba2**	
	Aa2			**Ba3**	Ongoing uncertainty
	Aa3			**B1**	High risk obligations
	A1	Strong payment capacity		**B2**	
	A2			**B3**	
	A3		**DEFAULT RATING**	**Caa**	Current vulnerability to default or in default
	Baa1	Adequate payment capacity		**Ca**	In bankruptcy or default.
	Baa2			**D**	
	Baa3				

Table 19: Standard & Poor's country risk ratings: average one-year transition rates (1975-1999)

	AAA	**AA**	**A**	**BBB**	**BB**	**B**	**CCC**	**SD**
AAA	97.45	2.55	0.00	0.00	0.00	0.00	0.00	0.00
AA	0.71	97.14	0.71	0.00	0.71	0.71	0.00	0.00
A	0.00	4.05	93.24	2.70	0.00	0.00	0.00	0.00
BBB	0.00	0.00	5.33	88.00	5.33	1.33	0.00	0.00
BB	0.00	0.00	0.00	7.06	83.53	7.06	0.00	2.35
B	0.00	0.00	0.00	0.00	14.81	81.48	0.00	3.70
CCC	0.00	0.00	0.00	0.00	0.00	33.33	33.33	33.33

Source: Standard & Poor (2000)

Table 20: Ratings

Countries	S&P ratings (1998)	S&P preorder (1998)	LRS ratings (1998)	Non-recursive regression ratings (1998)	S&P ratings (1999)	S&P preorder (1999)	LRS ratings (1999)	Non-recursive regression ratings (1999)	Moody's ratings (1998)	The Institutional Investor ratings (1998)
Argentina	BB	10	-0.2768	13.913	BB	10	-0.263	14.186	9	42.7
Australia	AA	19	-0.0289	20.117	AA+	20	-0.0128	21.369	19	74.3
Austria	AAA	21	-0.0094	19.655	AAA	21	0.0038	20.524	21	88.7
Belgium	AA+	20	-0.0476	17.282	AA+	20	-0.0439	18.443	20	83.5
Bolivia	BB-	9	-0.366	9.132	BB-	9	-0.3518	8.877	8	28
Brazil	BB-	9	-0.3744	9.164	B+	8	-0.4016	9.766	7	37.4
Canada	AA+	20	-0.0241	20.269	AA+	20	-0.0112	20.850	20	83
Chile	A-	15	-0.191	15.285	A-	15	-0.1841	15.155	14	61.8
China	BBB+	14	-0.2159	13.589	BBB	13	-0.224	13.371	15	57.2
Colombia	BBB-	12	-0.3854	8.337	BB+	11	-0.3964	8.738	12	44.5
Costa Rica	BB	10	-0.2748	11.586	BB	10	-0.257	12.418	11	38.4
Croatia	BBB-	12	-0.297	11.866	BBB-	12	-0.3202	11.360	12	39.03
Cyprus	A+	17	-0.1081	17.711	A	16	-0.1021	18.143	16	57.3
Czech Republic	A-	15	-0.2088	15.063	A-	15	-0.1904	15.436	14	59.7
Denmark	AA+	20	-0.048	20.373	AA+	20	-0.0492	20.729	20	84.7
Dominican Rep	B+	8	-0.3568	8.768	B+	8	-0.3431	9.650	10	28.1
Egypt	BBB-	12	-0.2915	11.247	BBB-	12	-0.3067	10.710	11	44.4
El Salvador	BB	10	-0.3379	10.460	BB+	11	-0.3301	10.482	12	31.2
Estonia	BBB+	14	-0.2518	12.486	BBB+	14	-0.245	12.602	14	42.8
Finland	AA	19	-0.064	19.362	AA+	20	-0.0458	20.005	21	82.2
France	AAA	21	-0.0828	18.262	AAA	21	-0.0614	19.179	21	90.8
Germany	AAA	21	-0.001	20.149	AAA	21	0.0126	20.959	21	92.5
Greece	BBB	13	-0.2255	13.239	A-	15	-0.1917	14.452	14	56.1
Hong-Kong	A	16	-0.017	18.360	A	16	0.0213	18.594	15	61.8
Hungary	BBB	13	-0.2442	12.623	BBB	13	-0.247	13.483	13	55.9
Iceland	A+	17	-0.047	20.544	A+	17	-0.0378	21.400	18	67
India	BB	10	-0.4063	8.548	BB	10	-0.3994	8.622	10	44.5
Indonesia	CCC+	5	-0.4576	4.821	CCC+	5	-0.4316	6.747	6	27.9
Ireland	AA+	20	-0.0179	18.929	AA+	20	-0.0249	20.491	21	81.8
Israel	A-	15	-0.2215	13.934	A-	15	-0.2189	14.200	15	54.3
Italy	AA	19	-0.1064	17.066	AA	19	-0.1122	17.362	18	79.1
Japan	AAA	21	-0.0604	19.106	AAA	21	-0.0506	20.079	20	86.5
Jordan	BB-	9	-0.323	10.532	BB-	9	-0.2818	10.574	9	37.3
Kazakhstan	B+	8	-0.4095	7.715	B+	8	-0.4048	9.005	9	27.9
Korea. Rep.	BB+	11	-0.2649	12.822	BBB	13	-0.2182	15.386	11	52.7

Latvia	BBB	13	-0.3026	11.281	BBB	13	-0.3039	11.316	13	38
Lebanon	BB-	9	-0.3223	10.625	BB-	9	-0.3121	9.342	8	31.9
Lithuania	BBB-	12	-0.3247	11.255	BBB-	12	-0.3233	10.621	11	36.1
Malaysia	BBB-	12	-0.1676	13.589	BBB	13	-0.1712	14.235	12	51
Malta	A+	17	-0.0999	16.302	A	16	-0.2402	16.131	15	61.7
Mexico	BB	10	-0.3608	9.548	BB	10	-0.3284	10.951	10	46
Morocco	BB	10	-0.2952	10.587	BB	10	-0.2881	10.198	11	43.2
Netherlands	AAA	21	0.0251	20.525	AAA	21	0.0337	21.009	21	91.7
New Zealand	AA+	20	0.0001	19.613	AA+	20	0.0001	19.661	19	73.1
Norway	AAA	21	0.0125	22.234	AAA	21	-0.0076	22.399	21	86.8
Pakistan	CC	2	-0.4563	5.184	B-	6	-0.4501	6.236	5	20.4
Panama	BB+	11	-0.2712	11.039	BB+	11	-0.2487	11.522	11	39.9
Paraguay	BB-	9	-0.3865	7.920	B	7	-0.4066	7.481	7	31.3
Peru	BB	10	-0.3536	10.386	BB	10	-0.3644	10.250	9	35
Philippines	BB+	11	-0.3242	9.595	BB+	11	-0.349	9.940	11	41.3
Poland	BBB-	12	-0.2772	12.718	BBB	13	-0.2743	12.899	12	56.7
Portugal	AA	19	-0.0742	17.556	AA	19	-0.0706	17.436	19	76.1
Romania	B-	6	-0.3987	8.000	B-	6	-0.3942	8.324	6	31.2
Russia	CCC-	3	-0.4428	5.161	SD	0	-0.4197	6.489	6	20
Singapore	AAA	21	0.0073	19.716	AAA	21	0.0225	18.543	20	81.3
Slovak Republic	BB+	11	-0.2814	11.386	BB+	11	-0.269	12.736	11	41.3
Slovenia	A	16	-0.1922	14.319	A	16	-0.1878	15.275	15	58.4
South Africa	BB+	11	-0.2523	10.920	BB+	11	-0.2386	12.873	12	45.8
Spain	AA	19	-0.0924	17.433	AA+	20	-0.0798	18.238	19	80.3
Sweden	AA+	20	0.0106	19.310	AA+	20	0.0143	19.538	19	79.7
Switzerland	AAA	21	0.071	23.436	AAA	21	0.0613	23.763	21	92.7
Thailand	BBB-	12	-0.2452	11.207	BBB-	12	-0.2383	13.127	11	46.9
Trinidad & Tob	BB+	11	-0.2824	11.624	BBB-	12	-0.248	12.681	11	43.3
Tunisia	BBB-	12	-0.2488	11.346	BBB-	12	-0.242	12.920	12	50.3
Turkey	B	7	-0.4458	6.388	B	7	-0.4177	6.407	8	36.9
UK	AAA	21	-0.0057	20.958	AAA	21	0.0062	20.769	21	90.2
United States	AAA	21	0.0205	23.005	AAA	21	0.0264	23.981	21	92.2
Uruguay	BBB-	12	-0.2695	12.516	BBB-	12	-0.2409	14.123	12	46.5
Venezuela	B+	8	-0.4444	6.999	B	7	-0.3921	8.422	7	34.4

Appendix: Correlation of difference matrices

Let us associate with any n-vector v its *[n x n]* "difference matrix" D whose component *(i,j)* is the difference of the *i's* and *j's* components of v. In Section 4.4.2, we have used that the correlation between any two n-vectors equals the correlation between their difference matrices. In spite of the elementary nature of this statement, we could not find it in the literature, and shall therefore provide here a formal proof of it.

Let us consider two n-dimensional vectors a and b, and two *[n x n]* dimensional difference matrices C and D, the elements of which c_{ij} and d_{ij} are given by:

$$\begin{cases} c_{ij} = a_i - a_j, \\ d_{ij} = b_i - b_j, \end{cases} \qquad i, j = 1, \ldots, n \tag{1.14}$$

It can be seen from (1.14) that:

- C and D are anti-symmetric:

$$c_{ij} = -c_{ji} \text{ and } d_{ij} = -c_{ji} \ , \quad i, j = 1, \ldots, n, \tag{1.15}$$

- C and D have diagonal elements equal to 0:

$$c_{ii} = d_{ii} = 0, \quad i = 1, \ldots, n \tag{1.16}$$

- the average of the elements of C and D is equal to 0:

$$\overline{c} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} / n^2 = 0 \ \ , \ \overline{d} = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} / n^2 = 0, \tag{1.17}$$

We shall show that the correlation between the vectors a and b,

$$\rho(a,b) = \frac{\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(b_i - \overline{b})^2}} \tag{1.18}$$

is equal to the correlation between the matrices C and D

$$\rho(C,D) = \frac{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(c_{ij} - \overline{c})(d_{ij} - \overline{d})}{\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(c_{ij} - \overline{c})^2}\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(d_{ij} - \overline{d})^2}} = \frac{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)(b_i - b_j)}{\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)^2}\sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(b_i - b_j)^2}} \tag{1.19}$$

The expression $\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)^2$ can be rewritten as

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}[(a_i - a_j)^2 + (a_j - a_i)^2] = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)^2 \tag{1.20}$$

Also, the expression $\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)(b_i - b_j)$ can be rewritten as

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}[(a_i - a_j)(b_i - b_j) + (a_j - a_i)(b_j - b_i)] = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)(b_i - b_j) \quad (1.21)$$

Using (1.20) and (1.21), the correlation between the matrices C and D (1.10) can be rewritten as :

$$\rho(C,D) = \frac{\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)(b_i - b_j)}{\sqrt{\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)^2}\sqrt{\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(b_i - b_j)^2}} \quad (1.22)$$

Similarly, the expression $\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})^2$ in (1.18) can be rewritten as

$$\frac{1}{n}\sum_{i=1}^{n}(a_i - \frac{1}{n}\sum_{j=1}^{n}a_j)(a_i - \overline{a}) = \frac{1}{n^2}\sum_{i=1}^{n}(\sum_{j=1}^{n}(a_i - a_j))(a_i - \overline{a}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)(a_i - \overline{a})$$
$$= \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}[(a_i - a_j)(a_i - \overline{a}) + (a_j - a_i)(a_j - \overline{a})] = \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)(a_i - \overline{a} - a_j + \overline{a}) = \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)^2 \quad (1.23),$$

while the expression $\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})(b_i - \overline{b})$ in (1.18) can be rewritten as

$$\frac{1}{n}\sum_{i=1}^{n}(a_i - \frac{1}{n}\sum_{j=1}^{n}a_j)(b_i - \overline{b}) = \frac{1}{n^2}\sum_{i=1}^{n}(\sum_{j=1}^{n}(a_i - a_j))(b_i - \overline{b}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i - a_j)(b_i - \overline{b})$$
$$= \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}[(a_i - a_j)(b_i - \overline{b}) + (a_j - a_i)(b_j - \overline{b})] = \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)(b_i - \overline{b} - b_j + \overline{b}) \quad (1.24)$$
$$= \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)(b_i - b_j)$$

Using (1.23) and (1.24), the correlation between the vectors a and b (1.19) can be rewritten as

$$\rho(a,b) = \frac{\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)(b_i - b_j)}{\sqrt{\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_i - a_j)^2}\sqrt{\frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(b_i - b_j)^2}}, \quad (1.25)$$

showing its equality to $\rho(C,D)$. QED.