EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

WORKING PAPER NO. 123

ANALYSING AND COMBINING MULTIPLE CREDIT ASSESSMENTS OF FINANCIAL INSTITUTIONS

BY EVANGELOS TABAKIS AND ANNA VINCI

February 2002

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

WORKING PAPER NO. 123

ANALYSING AND COMBINING MULTIPLE CREDIT ASSESSMENTS OF FINANCIAL INSTITUTIONS[1]

BY EVANGELOS TABAKIS AND ANNA VINCI[2]

February 2002

1 The paper has benefited substantially from the accumulated experience of Risk Management Division staff in the areas of credit risk monitoring and evaluation. In particular the authors would like to thank Philippe Molitor and Ken Nyholm for their valuable contributions and Carlos Bernadell for his continuous encouragement at all stages of the work The authors are grateful to a number of their ECB colleagues for comments received and discussions held on the contents of this paper. However, the views expressed in this paper are those of the authors alone and do not necessarily reflect views of the European Central Bank or the Eurosystem.

2 European Central Bank, Risk Management Division, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany.

Address	**Kaiserstrasse 29**
	D-60311 Frankfurt am Main
	Germany
Postal address	**Postfach 16 03 19**
	D-60066 Frankfurt am Main
	Germany
Telephone	**+49 69 1344 0**
Internet	**http://www.ecb.int**
Fax	**+49 69 1344 6000**
Telex	**411 144 ecb d**

ISSN 1561-0810

Contents

Abstract

The last consultative papers of the Basel Committee on Banking Supervision set the path for a future where a wealth of credit assessment sources may be available. New external credit assessment institutions and internal ratings-based assessments will be added to ratings of major international rating agencies and to benchmark assessment methods used by supervisors or central banks. In its first part, this paper contributes to the development of a toolbox to analyse and compare credit assessments by examining the ratings of three leading rating agencies on a set of credit institutions. The analysis decomposes the historical default rate corresponding to a rating into two components drawing on a "core" of published information and an "analyst contribution". In the second part of the paper, correlation and variance analysis of the analyst contributions lead to a combination of the available ratings, building on both the common core and the analyst part.

JEL classification: C21, C51, G15, G21, G28

Key words: ratings, default probabilities, analysis of variance, credit risk, analyst assessment, New Basel Capital Accord, Internal-Ratings Based approach, outliers, synthetic ratings.

Non-technical summary

The need to access and consistently use multiple credit assessments poses a number of difficult problems to the credit manager as regards the most appropriate tools for credit assessment and the most efficient way to combine them. In this respect, the consultative papers of the New Basel Capital Accord (Basel II) are a welcome addition for two reasons. Firstly, because they provide a set of guidelines that can be used to evaluate the quality of an assessment. Secondly, because they trigger new discussions on the challenges of credit assessment methodologies, their analysis, consistency and combination. Such credit assessments include but are by no means restricted to those provided by the leading international rating agencies, which have been the focus of attention of both the academic literature and the research departments of the agencies.

In this paper we develop tools to perform two different but closely related tasks:

1. Analyse credit risk assessments produced by different sources (e.g. leading rating agencies, smaller external credit risk assessment institutions and internal rating systems of commercial banks).

2. Combine multiple assessments to produce one single benchmark assessment.

The paper includes a review of the most recent academic work on external credit assessments and a summary of those points in the consultative papers of Basel II that could have a bearing on the availability and variability of credit risk assessment in the future. The central contribution of this paper is a model for the credit assessment process based on the decomposition of the default probability corresponding to a rating grade in a core part related to published financial data and an "analyst contribution". The analyst contribution is defined as the subjective interpretation of a bank's creditworthiness, which is often carried out by analysts and leads to a refinement of a rating. The model is applied on data collected for a set of 67 major banks, including all the ones in the EURIBOR panel, and using the ratings of Standard and Poor's, Moody's Investors Service and Fitch.

The first important result of the paper is that there is an advantage in analysing and combining assessments based on default probabilities that correspond to rating grades. While the three major international agencies do display statistically significant differences in rating credit institutions, there are no statistically significant differences in the historical default probabilities corresponding to these ratings. Hence, historical default rates naturally "correct" rating inconsistencies.

The data analysis, through a series of regressions, provides statistical evidence that agencies' ratings depend on balance sheet information, the country of incorporation and the bank's specialisation. Such information, however, is not enough to duplicate the agencies' methodologies and reproduce the agencies' ratings. The agencies' analysts are seen to deviate from this core rating by fine-tuning it based on additional information they may have, on their general rating guidelines, and on ad hoc assessment.

The proposed model produces a decomposition of the rating into a core, based on information available to everyone, and an analyst contribution, consisting of a more subjective interpretation of a bank's creditworthiness. In particular, additional information is extracted by measuring the impact of the analyst contribution and by analysing the correlation between such contributions by different agencies.

As a by-product of the analysis we also obtain a measure of convergence of different credit assessments. In the context of regulatory capital determination, this is a useful tool to reduce incentives for "credit rating shopping" by better pinpointing the ratings' range of variation and the causes of their divergence. Splitting ratings into a part based on easily available information and a part linked to the value added of rating agencies helps smooth the variance of credit risk assessments provided by multiple sources.

Finally, and although much work is yet to be done in the area of combining credit assessments, the paper puts forward the idea of using the analyst contributions to compute optimal combinations of different credit risk assessments, in order to derive one single benchmark or reference rating. Such combinations consist of adding a function of the analyst contributions to the core assessment, which is based on commonly used information. This function is selected in a way that reduces the variance of the analyst contribution and the influence of extreme assessments while, at the same time, using all the available information.

The paper wishes to provide useful tools both to supervisors and to credit risk managers. Supervisors could use those to evaluate the quality of credit assessments both under the standardised and the IRB approach, as the proposed method of combining assessments provides them with a way to create benchmark ratings based on multiple sources of information. Credit risk managers can also profit from the methods of the paper both in developing proprietary credit risk technology and in using multiple sources of credit risk assessment. Most importantly perhaps, the paper could contribute to increasing awareness of the need to carefully scrutinise and skilfully use credit risk information.

1. Introduction

Reliable and accurate credit assessments of financial institutions are a very important pillar for a stable financial system. The primary sources have been the leading rating agencies, whose assessments can serve as a benchmark for other methodological approaches. Both the academic literature and the rating agencies themselves have contributed to the debate about the theoretical framework of credit assessments and the role of rating agencies. Moreover, the proposal for the New Basel Capital Accord (from now on referred to as Basel II) at the beginning of 2001 has spurred a new round of discussions over the appropriate tools for credit assessment in connection with both the standardised and the internal ratings based approach.

This paper focuses on two aspects of this discussion:

a) Analysing assessments produced by different sources (e.g. leading rating agencies, smaller external credit assessment institutions and internal rating systems of commercial banks). The way proposed to accomplish such a task is to identify a common "core" across the different assessments and concentrate the analysis on deviations from this core.

b) Combining multiple assessments to produce one single benchmark assessment. This is a vital problem faced, for example, when assessments appear to be near or below certain important thresholds set by supervisors, central banks or counterparties of financial transactions. Basel II has touched on this issue, but this paper suggests that the problem needs further study.

The emphasis is on general methodological principles. In the choice of financial institutions to examine, the paper chooses to study mainly credit institutions important for the European financial system. The data was selected for the 57 banks whose dealings contribute to the derivation of the euro-area inter-bank offer rate (EURIBOR). In addition, a number of banks from the US and Japan that are major players on an international level were added. The National Bank of Greece was added to represent the 12th member of the European Monetary Union, as at the time of the analysis it was not represented in the EURIBOR composition. In total, 67 banks were used in the analysis. Concerning the ratings, the paper uses the three major agencies as sources: Standard and Poor's (from now on S & P), Moody's Investors Service (from now on Moody's) and Fitch (formerly Fitch IBCA, Duff & Phelps). Long term ratings are used throughout the analysis.

The paper is structured as follows: in Section 2, the most recent academic work on rating assessments is reviewed. Section 3 summarises the proposals of the consultative papers of Basel II on the subject of credit risk assessment, and highlights how this paper can contribute to using ratings from multiple sources when available. Section 4 presents a model for the credit assessment process based on a decomposition of the default probability in a core part from published financial data and an "analyst contribution". The analyst contribution is defined as the subjective interpretation of a bank's creditworthiness, which is often carried out by analysts and leads to a refinement of a rating. This process typically allows the agency to take additional, often non-quantifiable elements into account. Section 5 uses statistical methods to compare the ratings and corresponding default probabilities of the banks in the chosen sample. The model proposed in Section 4 is then fitted onto the sample of banks using regression methods, and the results are presented, paying particular attention to the outliers of the regressions. Section 6 proposes a new approach in four different variations to address the problem of combining multiple assessments and of producing a "benchmark" rating. Section 7 presents the papers' conclusions.

2. Literature review

Two strands of the literature are concerned with the rating process and the emergence of split ratings and are thus relevant in the context of this paper. The first is mainly of a descriptive nature and explores the relative use of quantitative and qualitative information in the rating process. The second uses historical data to analyse the cross-sectional differences of opinion between rating agencies and tries to link these differences to various quantitative and qualitative variables. This section reviews some of the most important contributions to these directions and relates these to the analysis presented in this paper.

A recent and quite detailed contribution to the first category in the literature is the paper by Grouchy, Galai and Mark (2001), which describes the approach used by the main credit rating agencies to produce credit ratings. The paper investigates the procedures applied by Moody's and S & P and explores the quantitative and qualitative variables that form the basis data on which the final credit score rests. The main objective of their study is to outline a "best practice" for designing internal rating systems in banks. However, they also touch on the aspects regarding the emergence of split rating. Without drawing any conclusions on the matter, the authors mention that smaller agencies, such as Duff and Phelps and Fitch (now merged), tend to assign higher ratings to the same debt issues than Moody's and S & P do. The authors mention a number of reasons for the prevalent differences, such as judgmental factors entering the evaluation process and differences in the rating methodology applied.[3]

In the context of the value of public information, Estrella, Park and Peristiani (1999) examine the predictive power of capital ratios (similar to the requirements of the Basel Accord) on bank failures in the US. They demonstrate that simple ratios predict failure as well as the more complex risk weighted ratios evaluated in the context of a logit model. What is perhaps more interesting in the context of the present study, is that, as the authors show, there is a close connection between the capital ratios and debt ratings published by rating agencies. Even though their data sample is limited, the results indicate that a significant part of debt ratings can be replicated by using published information such as balance sheet data.[4] Furthermore, Blume, Lim and MacKinlay (1998) show that accounting ratios and market-based risk measures are more informative in the rating of larger companies than in the rating of smaller companies. Some studies include *firm size* as an explanatory variable to control for the fact that larger companies tend to have more stable product lines and thus more stable financial ratios. Blume et al extend this argument

---

3 Indeed, S & P (1999) show that there is plenty of room for judgmental assessments entering the rating process.

4 Using a different data set and a different methodology Cisneros, Garcia and Sierra (p. 20, 1997) produce results in support of this conclusion, while Moody's investor service argues the opposite conclusion holds true.

and explicitly examine the relative informativeness of public information for larger and smaller companies. Using an order-probit model and cross-sectional data covering the period 1978 to 1995, they provide evidence in favour of the hypothesis that accounting ratios and market risk data are most informative for the group of larger companies. The result is derived for ratings in general, i.e. not specifically for the banking industry. However, if it can be extended to the ratings of banks, it implies that decomposing single ratings and evaluating the information content of the quantitative component vs. the judgmental analyst's component will have to be conducted depending on the size of the bank.

Extending the discussion of the informational issues touched upon by Blume et al, but using a proprietary data set consisting of ratings made by German banks, Brunner, Krahnen and Weber (2000) document that both public and non-public information play an important role in the rating process. They also conclude that substantial variation in the applied methodologies exists. Their main empirical result is that qualitative factors play an important role in the rating process and that they tend to improve the final rating value assigned to the company/bank being rated. Using a probit analysis, the authors are able to show that in one-third of the rating cases the qualitative information was decisive in terms of the final rating assigned. The papers by Tracy and Carey (1998) and Carey and Hrycay (2001) deal with the rating process and the estimation of default probabilities in internal credit risk systems of major US banks.

The second category of papers reviewed examines differences between ratings of the same company/bank published by different agencies. This strand of the literature provides a wealth of results demonstrating how and why ratings for the same issuer differ from agency to agency. In particular, a low level of agreement on ratings emerges when more than two major agencies are included in the comparison. As pointed out by BIS (2000) and Grouchy et al, one possible explanation for this result is selection bias, since most agencies, other than Moody's and S & P, only publish ratings on request, although Cantor and Packer (1997), who explicitly control for this effect in their analysis, refute this.

Ederington (1986) examines the ratings published by Moody's and S & P, finding a high degree of consensus in the way that accounting information is mapped into ratings. He concludes that split ratings occur for borrowers or issues that lie close to the borderline between two rating-scale values and that these split ratings are due to random noise rather than fundamental differences in the rating methodology applied by the rating agencies. In contrast to this result, Beattie and Searle (1992a) find that differences between several but not all pairs of rating agencies arise due to differences in the method of evaluation.[5] They use data from the Financial Times Credit Ratings International database to isolate pairs of rating agencies, which have co-existing ratings for more than 25 debt issues. These co-ratings form the basis of their analysis. In terms of general results they report that inter-rater agreement is reached only in 44% of the cases; 36% disagree by one category; 14% disagree by two categories; 4% disagree by three categories; and finally, 2% disagree by four or more categories. Interestingly, results are also shown for individual pairs: for example, S & P and Moody's disagree on 36% of the ratings. Likewise, S & P and

[5] It is important to bear in mind that the analysis of Beattie and Searle (1992a) covers several different rating agencies, where as the Ederington (1986) study focuses exclusively on Moody's and S & P. Hence, it is not surprising that the two articles reach different conclusions on the consensus between agencies.

Fitch disagree on 49% of the ratings, and in these cases, Fitch produces a higher rating in 69% of the cases; Moody's and Fitch disagree on 53% of the ratings, with Fitch producing higher ratings 65% of the times.

Beattie and Searle (1992b) extend the previous results by examining the reasons for the prevailing differences between the rating agencies. They find that ratings published for the banking and the utilities sectors exhibit a lower degree of consensus than that found within other sectors. Furthermore, they show that rating agencies suffer from a "home-country bias", i.e. systematically assigning higher ratings to banks and other companies located in their home country. In the same vein, Cantor and Packer (1994) analyse differences in international ratings of senior debt in banks and reach similar conclusions. Looking at 1,018 worldwide bank-ratings published by nine of the leading rating agencies and available as of January 1994, the authors show that the degree of consensus is low. Equality of ratings varies across the pairs of agencies from around 10% to 44%, which is comparable to the findings of Beattie and Searle (1992b). However, Cantor and Packer (1994) do not go into depth with respect to the reasons for the apparent differences in the ratings provided by the rating agencies, but they suggest that the main reason for split ratings is that the agencies apply different methodologies in the process of producing ratings.

The present study contributes to the literature in mainly two areas. First, it focuses on the common characteristics in the methodologies of major agencies and presents empirical results on the relative importance of the "core" and "analyst contribution" parts of the ratings published. This decomposition of ratings contributes to the discussion in the literature of why differences occur and whether they are mainly due to different methodologies being applied by the rating agencies or whether it is the judgmental part of the rating process that prevails. Second, the study demonstrates how to combine ratings from different rating agencies into a "rating benchmark", which can circumvent the difficulty of choosing between different ratings for the same asset or issuer.

3. Impact of Basel II

Basel II aims at better aligning regulatory capital charges to the risks inherent in the assets of a bank's balance sheet, typically market, credit, operational and liquidity risks. Risk weights, varying depending on the category of borrower, are used to map credit assessments into risk-weighted assets. Summed over all categories of borrowers, this provides the information on which the amount of regulatory capital is determined. The problems of analysing and combining credit assessments discussed in this paper are closely related to the proposals of Basel II and the difficulties envisaged in its implementation.

The new proposal foresees two sources of information on credit risk to determine the risk weights: external credit ratings and internal assessments. Both should be treated on an equal footing and hence need to be handled consistently in order to secure the quality of the information. This consistency ought to be achieved not only across agencies, but also across issuer categories (e.g. corporate versus sovereign) and over time. Accordingly, for the measurement of credit risk inherent in different assets, banks can adopt either a standardised or an internal rating based (IRB) approach. The former relies on agencies' ratings, while the latter relies on a bank's internal assessment of its counterparties and exposures, and foresees a two-step implementation process, namely a 'foundation' and an 'advanced' phase.

3.1 Credit risk information from external credit rating institutions

Under the standardised approach, the weights for determining risk-weighted assets in the banking books are based on credit assessments provided by External Credit Assessment Institution (ECAI). In order to be eligible, ECAIs undergo a supervisory recognition process. This process determines whether they comply with a set of six eligibility criteria: objectivity, independence, international access and transparency, disclosure of information, sufficient resources and finally credibility.

National banking supervisors will hence play an important role in such validation process, and there is widespread concern on how consistently supervisors will exercise their judgement in applying and interpreting the criteria across countries. Indeed, the analysis in Section 5 of this paper shows that credit assessments of financial institutions by the major international rating agencies are considerably influenced by the country of incorporation of the institution.

To guarantee that the recognition process be open, fair and consistently applied across countries, the Basel Committee on Banking Supervision (BCBS) is called to take an active role in formulating specific recognition criteria. In this context the idea of a "college of supervisors" established under the BCBS to help in the implementation of consistent ECAI eligibility, mapping and consistency of the IRB approach across countries is very prominent. The list of recognised ECAIs should consequently be made public. Sections 4 and 5 provide tools that could help national supervisors to perform their evaluation in a consistent manner.

There is a second important area of responsibility of supervisors: the slotting of ECAIs' assessments into the standardised risk-weighting framework in order to allow the different ratings to be compared on a normalised scale. As banks are asked to choose an ECAI and to use its rating consistently for each type of asset, it is important to prevent "shopping" for credit assessments provided by different ECAIs. Such cherry-picking would lead banks to use the highest available rating to determine their respective level of regulatory capital. As banks might still be forced to fall back on two or more ECAIs due to the non-availability of ratings for different assets, Basel II proposes a rule in order to choose the appropriate risk weight. Section 6 of this paper addresses the problem of optimally combining ratings by comparing different methodologies.

3.2 Credit risk information from internal assessments

If banks choose an IRB approach, they must provide internal estimates of borrower creditworthiness to evaluate the credit risk inherent in their balance sheet's assets. Both at the outset and on an ongoing basis, the IRB approach remains valid if it fulfils a number of minimum requirements. These are aimed at guaranteeing the integrity and credibility of a bank's rating system as well as ensuring that the calculation of regulatory capital and the estimation of the risk components are correct and reliable.[6]

The categories of minimum requirements concern those key elements of the bank's internal rating and risk measurement process. Each category has a vital role to play in determining the overall quality of the information that banks would provide to supervisors for determining minimum capital requirements. In particular, Basel II endorses features that promote meaningful identification and differentiation of estimated borrower risk across exposures, reliable and disciplined estimation of risk components, and clarity in the documentation of rating systems and decisions. Interestingly, the Basel Commission specifically states that when validating adherence to these minimum requirements, supervisors will need to make objective comparisons as well as subjective judgements. The minimum requirements thus include objective and measurable criteria as well as more subjective, judgement-oriented criteria. The same principle is followed in the analysis of credit assessments presented in sections 4 and 5 of this paper.

A fundamental difference distinguishes the IRB approach from relying on external credit assessments in the context of regulatory capital determination. Whereas ECAIs need to publish their service or product (i.e. the credit rating for a specific asset or borrower), banks that adopt the IRB approach only have to publish the nominal amounts related to the different risk components for each portfolio. No obligation exists to publicly release the primary risk component, namely the probability of default (PD) related to a specific asset or borrower, which constitutes the "internal" credit rating. Similarly, Basel II does not require specific risk rating definitions, not does it assign specific levels of risk estimates that must be associated with each internal grade. The differentiation not only sets apart external (ECAI) or internal (IRB) credit assessments. It also highlights that whereas credit information is a service or product in the former case, in the latter it is one outcome or phase of a bank-wide credit risk assessment process.

3.3 How to make use of the availability of multiple credit risk assessments

The comparability of external and internal ratings is currently not addressed by the regulator. Although, as mentioned above, supervisors would map ECAIs' assessments into the standardised risk weighting framework and hence provide a normalised scale of the different ratings, such mapping is neither foreseen nor required by Basel II for the IRB approach.

Although ratings supply an evaluation of a company's credit riskiness, they generally lag behind market developments. Agencies explain this lag with the reliance on accounting and financial statements, with the impossibility of continuous monitoring and with a certain reluctance to be criticised for rating volatility, which is also very much related to rating agencies complying with the credibility criterion as indicated in Basel II.

---

6 The minimum requirements for the foundation IRB approach are described in detail in chapter 2 of the Basel Committee's consultative document on "The Internal Ratings-Based Approach".

The statistical evaluation of different credit risk assessments and the proposed ways to compare and combine them presented in the next sections could be used in at least three respects in the context of a selection or validation procedure.

1. Firstly, in the case of ratings provided by ECAIs, the analysis presented in this paper allows risk managers to shed light on the structure of the rating process by comparing the assessment under consideration with basic, commonly available information. The analysis shows how to examine whether a certain rating lies within an acceptable range, and pinpoints cases of unusual deviations from this range.
2. Secondly, it allows the contributions of the analyst to be evaluated by examining correlations with other analyst assessments. In the case of an internal rating model under the IRB approach that has the same risk grades as the ones used currently by the major ECAIs, the analysis in this paper may also provide a benchmark to examine the validity of ratings produced by an internal model.
3. Thirdly, the analysis can be helpful for the mapping procedure of banking supervisors and for the establishment of a normalised rating scale. These two elements are an important part of banking supervision that foster comparability and qualitative improvements of the credit assessments produced. The paper provides empirical evidence on the value of default probabilities as a tool in the mapping procedure (Section 5.1) and proposes ways to combine different assessment (Section 6) in one rating scale.

The next sections propose a new way to model credit risk assessments and alternatives to compare and combine them.

4. Modelling credit assessments

This section presents a model explaining the "rating behaviour" of an agency. The main goal is to decompose the rating of an agency into a component based on published and quantifiable information and a component attributed to additional information or to the subjective view of the agency's analyst. Here and in the analysis to follow, the analyst contribution is defined as the subjective interpretation of a bank's creditworthiness, which is often carried out by analysts and leads to a refinement of the core rating. As in much of the literature,[7] the analysis assumes that agencies or risk managers produce a basic rating or score by using quantifiable information. Building on this, the credit risk analyst then refines the core evaluation with a subjective evaluation of the financial results, of the company's growth potential, of announcements, or of possible mergers.

An innovative feature of the model is that it relies on the default probabilities that correspond to rating grades rather than rating grades themselves. The advantages of using default probabilities are explained in detail by the empirical results of Section 5.1. Let p_i be the (unknown) default probability of bank i. Rating agency j produces a rating for this bank corresponding to an estimated default probability p_i^j. A monotone

[7] For example, Treacy & Carey (1998).

transformation 'f' is applied on these probabilities to transform them from a number in the interval [0,1] to a real number. While this step may seem a technical detail, it makes it possible to use linear regression methods. The two most commonly used transformations for this purpose are the logit transformation:

$$f(p) = \ln \frac{p}{1-p}$$

and the probit transformation:

$$f(p) = \Phi^{-1}(p), \text{ where } \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

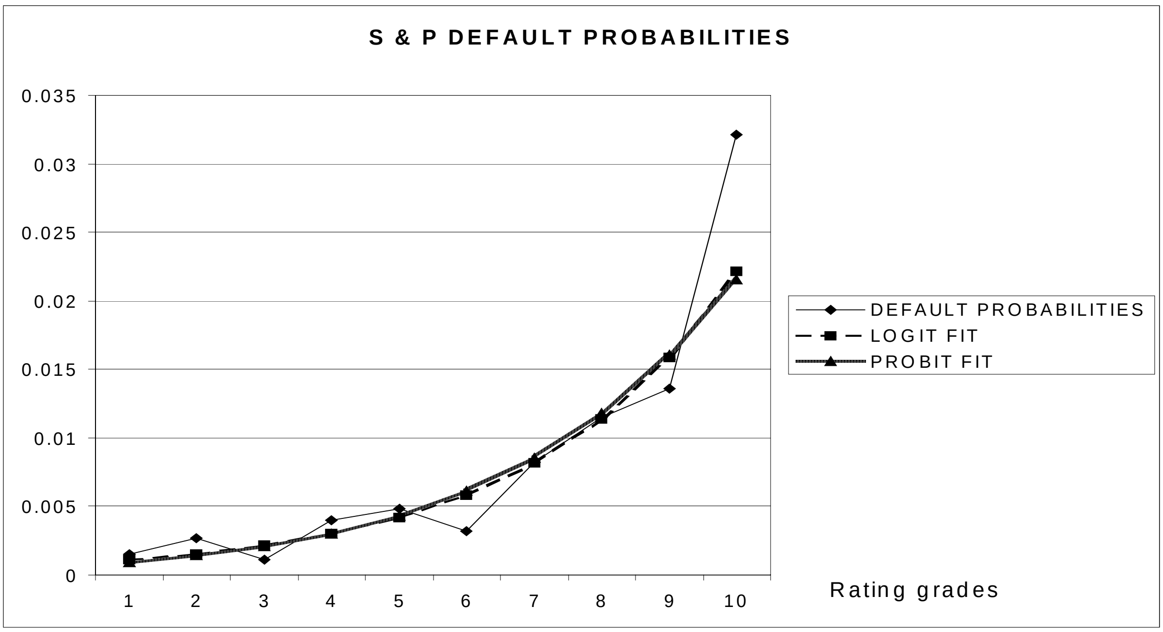
These transformations are supported by the assumption that default probabilities can be derived by computing a rating score 's' for each rated bank, which then determines the default probability according to either the logit or normal distribution:

$$p = \Pr(default \mid score = s) = \frac{e^{-s}}{1+e^{-s}}, \quad \text{or} \quad p = \Pr(default \mid score = s) = \Phi(s).$$

Applying the above mentioned transformations 'f' recovers the score 's' for every bank.

A technical problem arises from the fact that the historical default rates for each rating grade are not an increasing function of the rating numerical equivalent (see Annex 2). It is therefore necessary to smooth the default probabilities by fitting either a logit or a probit distribution to the historical default probabilities (see figure below). The chart also shows that there is essentially no difference between a logit and a probit fit for the range of rating grades 1 to 10. These "smoothed" default rates are used also in the rest of the paper as estimates of the default probabilities corresponding to the various rating grades. The logit transformation is used throughout. Furthermore, no distinction is made between the rating of the agency and the corresponding estimate of the default probability that this rating implies, unless such a distinction is explicitly mentioned.

Chart 1 – S&P Default probability and the associated smooth logit and probit fits



In summary, the steps leading from a rating's numerical equivalent in the integer scale of 1 to 10 to a transformed smoothed default probability that can be used as the dependent variable in a linear regression are outlined below:

1. The rating is mapped to the historical default probability corresponding to this rating grade for the agency in question.
2. The default probabilities are smoothed, using a logit transformation, to become a monotone function of the rating. To avoid introducing additional notation, p_i^j is also used for the smoothed default probability assigned by agency j to bank i.
3. The logit transformation is applied to the smoothed probabilities to recover the "score" $f(p_i^j)$ assigned by agency j to bank i.
4. The obtained score (transformed smoothed default probability) becomes the basis for further analysis.

It is now assumed that the agency's estimate of the default probability is subject to a random error satisfying:

$$f(p_i^j) = f(p_i) + \varepsilon_i^j \qquad \textbf{(1)}$$

In the above equation the index j=1,2, …, J identifies the J different agencies and the index i=1,2,…,I identifies the I different banks. In providing the rating of bank i, agency j makes use of a vector X_i of publicly available information. We will refer to this information as the **"core" information** in the understanding that it does not contain all the publicly available information but rather a number of easily available and readily quantifiable variables which are expected to be used by an assessment system in producing bank ratings. The rating agency makes use of this information in a "linear way" which is translated into the assumption:

$$E(f(p_i^j) \mid X_i) = (a^j)' X_i \quad \textbf{(2)}$$

where a^j is a vector of appropriate dimension. Written in an alternative way, the rating of agency j for bank i can be decomposed (after the appropriate transformation) as follows:

$$f(p_i^j) = (a^j)' X_i + r_i^j \quad \textbf{(3)}$$

where the residual r_i^j represents the **contribution of the analyst** to the final rating.[8] This analyst contribution can include the processing of news and the outcome of analysis on additional variables that, although possibly also publicly available, may not be used by all agencies, may be difficult to quantify, or

[8] A different but equivalent way to interpret the model of equation (3) is to think of it as an example of a single-layer neural network. Such networks are also known as perceptrons based on the original biological applications of their architecture. See, e.g. Bishop (1995). The network would use the vector *X* as input and provide the default probability as its output. Given the simple form of the network, its 'training' does not require an iterative method and the "weights" given by vector *a* can then be estimated using multiple linear regression.

their relevance for the rating process can be a matter of subjective assessment. Equations (2) and (3) imply that:

$$E(r_i^j \mid X_i) = 0 \text{ and } Cov(r_i^j, X_i) = 0 \quad \textbf{(4)}$$

These equations can be translated into the assumption of an "unbiased analyst", given the available information, and of lack of correlation between public information and the contribution of the analyst.

The agency's error can be decomposed as follows (using equations (1), (2) and (3)):

$$\varepsilon_i^j = f(p_i^j) - f(p_i) = E(f(p_i^j) \mid X_i) + r_i^j - f(p_i) = E(f(p_i) \mid X_i) + E(\varepsilon_i^j \mid X_i) + r_i^j - f(p_i) \quad \textbf{(5)}$$

Note that the part $q_i = E(f(p_i) \mid X_i) - f(p_i)$ is the same for every agency j. Therefore, comparisons between agencies can be based on $E(\varepsilon_i^j \mid X_i) + r_i^j$. As an example, the difference in the estimation error between agencies 1 and 2 for the same institution '*i*' will be:

$$\varepsilon_i^1 - \varepsilon_i^2 = E(\varepsilon_i^1 \mid X_i) + r_i^1 - E(\varepsilon_i^2 \mid X_i) - r_i^2 \quad \textbf{(6)}$$

We now assume that the agency provides unbiased ratings given the public information available, and therefore the error in the assessment is, on average, equal to 0. Concretely, it is assumed that:

$$E(\varepsilon_i^j \mid X_i) = 0 \Rightarrow \varepsilon_i^j = r_i^j + q_i \quad \textbf{(7)}$$

Thus, although the agency's error is not directly observable (because the true default probability p_i is not known), this error is equal to the analyst's contribution plus a term not depending on the agency. Comparisons between ratings can be thus based on estimates of the analyst's contribution r_i^j. In the example of equation (6), the difference in estimation error of the two agencies is reduced to:

$$\varepsilon_i^1 - \varepsilon_i^2 = r_i^1 - r_i^2 \quad \textbf{(8)}$$

Finally, in this case,

$$E(\varepsilon_i) = E(E(\varepsilon_i^j \mid X_i)) = 0 \quad \textbf{(9)}$$

and the agency's ratings are also unconditionally unbiased.

Note that this assumption implies that two different unbiased agencies j and k should produce, on the average, identical estimates of default probabilities since:

$$E(f(p_i^j) - f(p_i^k)) = E(\varepsilon_i^j - \varepsilon_i^k) = 0 \quad \textbf{(10)}$$

This is a hypothesis that can be tested empirically on a set of banks by comparing the ratings of different agencies for these banks and the corresponding estimates of default probabilities. The paired t-test or the non-parametric paired sign test can be used for this purpose, as shown in Section 5.1. A result showing insignificant differences between agencies makes it less probable that a systematic bias exists, although it does not definitely prove that the three agencies are unbiased. In fact, if no significant difference is detected, then a bias could only exist if it were exactly the same (in sign and size) for all three agencies. Such an event has to be considered highly unlikely for the case of three independent agencies. Clearly, a test of unbiasedness will always be necessary before the analysis of the next section can be applied to a new rating agency. However, we consider that such a test should be part of the assessment of any new agency by supervisors, since the existence of a systematic bias in the estimation of default probabilities should in principle disqualify the assessments of the agency. Therefore, the assumption that the agency's assessments are unbiased is considered a natural one.

5. Empirical evidence on S & P, Moody's and Fitch ratings.

This Section applies regression methods to compare the ratings and default probabilities of the selected sample of rated banks (Section 5.1) fitting the model presented in Section 4 (Sections 5.2 and 5.3). A set of exceptions is further examined in Section 5.4, focusing on the analyst component.

5.1 Comparing ratings and corresponding default probabilities

When comparing different credit assessments of the same financial institution by different rating agencies, there are two ways to proceed:

1. One can rely on a well established and agreed upon equivalence between rating grades of the different assessment institutions. Such a correspondence exists for the three rating agencies examined in this paper, as shown in Annex 2 for Moody's and S & P, and it reduces ratings to a simple numerical equivalent (e.g. AAA=Aaa=1, AA+=Aa1=2, etc.).
2. One can use estimates of the default probability corresponding to each rating grade of each agency. Such estimates can be the result of available studies on historical default frequencies for the different rating grades of the agencies. Studies are available for S & P and Moody's ratings (see, for example, BIS (2000) for the one- and five-year horizon default rates). In this approach, what is actually compared is not the rating but the corresponding default probability.

In terms of statistical methodology, the comparison can be performed by paired sample tests. The tests examine the differences between pairs of ratings (or estimates of default probability) for the same financial institution. To avoid dependence on the distribution of the numerical equivalent or the default probability, the tests used here were non-parametric: the paired sign-test and the paired rank test. The tests gave identical results in all paired comparisons. The results were rather surprising as detailed below:

- When comparing the rating numerical equivalent, differences between Fitch and Moody's did not prove statistically significant. On the contrary, S & P ratings were statistically lower than those of Moody's and Fitch.
- When comparing the estimated default probabilities, however, no significant differences were found between agencies.[9]

[9] Given the fact that no statistically significant difference between the ratings of Fitch and Moody's was found, the default rates of Moody's were also used for Fitch.

The most plausible explanation of these results is that the rating correspondence shown in Annex 2 may not be completely correct. In fact, when comparing the five-year default rates in S & P's AAA grade with those of Moody's Aaa, it can be seen that there is a significant difference (0.15% against 0.22%). Therefore, it is to be expected that S & P's AAA is in fact, a better rating than Moody's Aaa. Treating them as equivalent results in a "downgrade bias" for S & P. In this sense, mapping ratings into default probabilities eliminates this bias and is the preferred approach followed in this paper.

For the purposes of this paper, the above result is very important because it validates the use of the available historical estimates of default probabilities by rating grade instead of the grades themselves in the analysis to follow. In addition, the result is helpful because it highlights the importance of default probabilities in comparing credit assessments.

The result suggests that reliable estimates for default probabilities corresponding to the different rating grades can be used to map rating grades to risk weights. As it provides a common denominator for the mapping process of different assessments, such a technique removes the complexity of mapping different ratings to risk weights that is necessary to implement the standardised approach, as already flagged in the consultative document of Basel II.

An intrinsic shortcoming of the method presented in this paper occurs when no such default probabilities are available, although historical estimates of default probabilities are not the only possible source of estimates. Should historical data be unavailable, ex ante estimates based on a model of default developed by the ECAI or being part of an IRB system could be used instead.

5.2 Choosing appropriate financial ratios

To apply the model of the previous section, a vector of public information X must be chosen. It is important to note that no claim is made that the approach chosen in this analysis is indeed the one used by the rating agencies. In fact, it is not the aim of this section to duplicate the analysis of the agencies.[10] Instead, based on statistical evidence alone, a body of publicly and easily available information is chosen in order to link and compare it to the assessments of the agencies. The documentation of the agencies' research departments and a review of the literature were used to provide guidance as to the type of variables that could prove useful for the purposes of this analysis.

It is inevitable that any concrete analysis is restricted to a limited core subset X of the public information available. What may appear to be the result of the analyst contribution in the regressions presented in this section, could easily be explained in the core part if the information set X were to be augmented. The main source of financial data was BankScope. Theirs is also the definition and calculation of industry ratios, such as those used below. In addition and as a complement to BankScope, balance sheet data and information on financial ratios was gathered from annual reports relative to the

[10] At this stage of the analysis, information on ratings (or, equivalently, default probabilities) is not yet used, only the data on the 11 financial ratios, which is decomposed into three independent contributions.

year 2000. The following eleven ratios were chosen to represent the categories of profitability, efficiency, liquidity and capital structure:

- The Net Interest Margin, defined as the net interest after tax on total operating income;
- The Pre-tax Operating income on average assets;
- The Return on Average Assets (ROAA);
- The Return on Average Equity (ROAE);
- The Tier 1 ratio;
- The Total Capital ratio;
- The Equity on Total Assets;
- The Net loans on total assets;
- The Equity on Customer & Short Term Funding;
- The ratio of the total deposit portfolio to total assets, also referred to as 'Percentage of deposits lent';
- The Cost-to-Income ratio, defined as the operating costs on total operating income.

Based on theoretical considerations alone, all the above financial ratios should contribute to the creditworthiness profile of a financial institution. However, from an empirical point of view, a reduction of the number of factors to be considered becomes necessary for at least two reasons. Firstly, several of the above mentioned ratios can be expected to be strongly correlated with each other. For example, 'tier 1 ratio' and 'total capital ratio' were highly correlated. Likewise, the 'net interest margin' was positively correlated with 'equity on total assets', and 'pre-tax Operating income on average assets' with 'ROAA'. Secondly, data on all the above ratios was not available for all credit institutions considered in this analysis. In such cases, the analysis had to rely on the ratios for which the information was more complete.

To optimise the number of factors to be considered and eventually kept in the analysis to follow, a principal component analysis was performed on the 11 factors mentioned above. The analysis revealed that the information contained in the 11 variables can be summarised and explained by the following three components, provided below in order of importance:

1. One component measuring **profitability and efficiency**, highly correlated with 'Pre-tax Operating Income on Average Assets', 'Return on Average Assets', 'Return on Average Equity', 'Net Interest Margin' and 'Equity on Total Assets'.
2. One component measuring **capital adequacy and liquidity**, highly correlated with 'Tier 1 ratio' and 'Total capital ratio'.
3. One component representing 'net **loans** on total assets'.

This was also confirmed by an examination of the correlations between the 11 variables (see Annex 3).

Finally, it was considered helpful to add one measure of "**size**" of the bank as a possible factor. This reflects the information provided by the rating agencies on this issue and acknowledges a certain dependence of the rating on the banks size. Therefore, total assets were chosen to introduce this dimension. In fact, to produce a variable more symmetrically distributed around its median, the logarithm of total assets was chosen for further consideration.[11]

5.3 Decomposing ratings

Guided by the principal component analysis, regressions of the transformed default probabilities of the three agencies on subsets of the above 11 variables were performed. The selection of the subsets aimed at covering the three groups of financial ratios identified in the previous section as well as the factor "size". However, it became obvious from the results of the first regressions and was confirmed by the literature and the research papers of rating agencies that the country of incorporation and the bank's specialisation should also be taken into account. In order to do that, indicator (dummy) variables were introduced to include this information. In a second round of regressions, the regressors were chosen from an extended set that included, in addition to the 11 financial ratios and the variable 'total assets', dummy variables indicating the country of incorporation and the specialisation of each bank in the sample. The most successful combination of variables in terms of explanatory power appeared to be the following:

1) **Total assets (log-transformed)**

2) **Net loans on total assets**

3) **Equity on total assets**

4) **Japanese incorporation (dummy variable)**

5) **Italian incorporation (dummy variable)**

6) **Governmental support (German Landesbanken)[12] (dummy variable)**

The results of the regression of the (logit-transformed and smoothed) default probabilities on these variables are shown in Annex 4. The results indicate considerable agreement on both the choice and the significance of the above variables across all three rating agencies, although no claim is made here as to the actual use of these particular variables by the rating agencies. The only statistically valid result is the discovery of significant correlation between ratings (equivalently default probability estimates) and linear combinations of the above variables. This statistical fact is not refuted by the descriptions of the rating methodologies of the agencies by their own research departments although the agencies refrain from confirming that balance sheets constitute the core element of their methodology (Moody's Investor Services (2001)).

---

[11] An examination of the histogram of counterparties' total assets revealed a skewed distribution. Taking the logarithm of total assets resulted in a distribution closely approximating the normal.

[12] Note, however, that governmental support to a credit institution was not always explicitly declared in the data source.

It should be emphasised that the analysis performed here is not a forecasting exercise. The goal is not to forecast an assessment of a rating agency in advance of its announcement because there is no time lag assumed between the rating and the information in the independent variables used so that all regressions are performed in a cross-sectional setting. The exercise is rather one of **analysis of variance** where the observed variability in ratings between EURIBOR banks is attributed to either a set of basic core variables or to an analyst contribution. The question of robustness of the analysis (and the resulting conclusions) is addressed by a separate application of the method on the three agencies. We consider the fact that all three regressions revealed the same set of significant dependent variables (out of all possible combinations) and approximately the same explanatory power to be a strong indication of the method's robustness.

The results of the analysis on the three sets of residuals (analyst contribution) are shown in Annex 5. It is immediately obvious that S & P's residuals are larger (almost 1.5 times) than those of Moody's and Fitch. This would indicate that the analyst contribution is higher in S & P compared to the other two agencies.

Correlations between the three sets of residuals are high (about 0.85), pointing to the fact that, in general, analysts deviate in the same direction from the core information. This would also point towards a common approach in the rating methodology of the agencies, although S & P analysts appear to react more aggressively to the pool of information they have available and are more willing to deviate from a strict "scoring" approach to rating. We will rely on this interpretation of correlation to build synthetic ratings in Section 6.

As expected, explicit governmental support in the form of government guarantees has a significant effect on the rating. Furthermore, as apparent in the examination of individual banks in the next subsection, various forms of less explicit support also have a similar effect.

The correlation between the size of the institution and rating has always been a difficult issue to explore. A significant contribution of "total assets" to the rating could only be found after controlling for the Japanese banks (large but not as highly rated)[13] and the German Landesbanken (smaller but highly rated because of government guarantees).

Finally, in general rating agencies seem to consider that Italian banks display lower than average profitability margins and net interest revenues. In addition, the largest Italian banks are considered still quite small by international standards. Therefore, ratings of Italian banks tend to be lower than comparable banks incorporated in other countries of the European Union.

---

[13] Japanese banks are rated significantly below what their balance sheet date would indicate. After the financial crisis of the 90s, the safety net created with the establishment of public funds for banks with non-performing loans in their balance sheets reassured rating agencies that the government was committed to supporting the banking sector. However, authorities' intervention was applied reluctantly, mainly in an attempt to reassure the international financial community that the banking sector is not in need of further intervention and that market forces can operate independently. Concerns over downside risks have induced two global rating agencies, Moody's and Fitch, to downgrade Japan's sovereign rating in 2000. All indications are that agencies will continue to be extremely conservative in assigning ratings, especially as banks continue to struggle with low profitability and weak capital positions induced by bad corporate governance and structural problems, and as government support tends to diminish over time.

5.4 A study of outliers

The regressions identify a group of outliers, i.e. banks that are rated considerably better or worse than what the values of the regressors would actually predict. Table A lists the outliers and the "actual minus predicted" logit-transformed default probability. The column 'Average' shows the outliers when the simple average of the logit-transformed default probabilities from the three agencies are regressed, while the last three refer to the regression performed on the logit-transformed default probabilities of the three individual agencies in turn. Owing to the range of the logit scale (all negative numbers and increasing in absolute value with ratings), a positive (negative) difference in the table means that the predicted rating is higher (lower) than the actual one, according to the regression equation used in the analysis.

Table A - Outliers: banks names and residual difference between the actual and the predicted logit

Name of the bank and relative position in the list of Annex 1		AVERAGE	Fitch	S&P	MOODY'S
5	Banca di Roma		0.58		
9	Bank Austria AG - IAS			-0.77	
16	Banque et Caisse d'Epargne de l'Etat Luxembourg	-0.64			
21	CDC IXIS Capital Markets	-0.56	-0.51		-0.55
29	Crédit Industriel et Commercial - CIC				0.57
30	Crédit Lyonnais			0.91	
35	Deutsche Genossenschaftsbank DG BANK - IAS				0.60
45	Landesbank Baden-Wuerttemberg	-0.69	-0.59	-0.87	-0.58
52	National Bank of Greece SA	0.87	0.67	1.20	0.75
55	Rabobank Group	-0.78	-0.66	-1.05	-0.61

The comparison of actual to predicted default probabilities is not meant as a criticism to the agencies' ratings, nor does it imply that these ratings are false. Instead, this part of the analysis is to be perceived as a useful tool in understanding the reasons behind this apparent deviation and detect factors that played a role in the derivation of the rating but were not included in the regression.

If one rationalises the differences between actual and predicted logit-transformed default probability based on the specifics of each bank, three reasons for banks to be 'outliers' would emerge.

- Firstly, the status of a bank as regards **government guarantees** improves its credit quality, other things being equal. While the status or type of a bank has been taken into account in the regressions by including appropriate dummy variables (e.g. Landesbanken or securities house), other implicit or explicit guarantees provided to commercial or savings banks have not. In at least four cases highlighted by this research, the support of a public or administrative institution or a special regime that ensures maximum security of funds can be linked to the status of outlier. This information is difficult to quantify, and it remains up to the analyst to assess it and incorporate in the rating. The existence of guarantees is thus a substantial component of the analyst contribution.
- Secondly, the **geographical location** of a counterparty can induce an automatic re-scaling of ratings, such as in the case of Greek and Italian banks or, outside Europe, of Japanese banks. Recognising this fact, dummy variables were introduced for the case of Italian and Japanese banks in the regressions. The same could not be done with the only Greek bank in the sample, as this would artificially reduce the

residual to 0. Therefore, this bank appears as an outlier. The influence of the geographical location in the re-scaling of the rating can be considered the second dimension in the analyst contribution.

- Lastly, agencies are often reluctant to change ratings if **specific events** protracted in time affect an institution's creditworthiness, such as, for example, restructuring, mergers or announcements of strategic decisions or of future financial results. Agencies rather wait until the event's outcome is certain. In some other cases, the events are reflected in the ratings but not in the last available balance sheet data. Both situations lead to discrepancies between the predicted and the actual assessment. Hence, the choice to incorporate in the assessment recent events that can have a bearing on the bank's creditworthiness provides a third dimension of the analyst's contribution.

6. Comparing and combining ratings

Viewed over a long time-horizon it is unavoidable that rating agencies produce different ratings for the same company. In fact, in a broad economic context it can even be desirable that such differences emerge if they promote an increased understanding of the underlying financial situation of the bank/company being scrutinised. However, from a regulatory viewpoint, differences between ratings are problematic since they call for a conscious choice as to which of the prevailing rating values is valid. Also, should serious differences prevail for too long, doubts about the objectiveness of the whole rating concept could emerge.

6.1 The problem

Suppose that a number of ratings are available for a group of banks. As before (see Section 4), we continue to assume that these ratings can be mapped into default probabilities and that a monotone scale transformation f is applied to these probabilities. It is also assumed that these agencies had access to the same publicly available information X_i for bank i, and the transformed default probabilities are regressed on X_i:

$$f(p_i^j) = (\hat{a}^j)' X_i + \hat{r}_i^j$$

As before, the index j=1,2, …, J identifies the J different agencies and the index i=1,2,…,I identifies the I different banks. Let S denote the estimated covariance matrix of the residuals r. This is going to be a JxJ matrix consisting of the variances of the residuals for every agency in the diagonal and the covariances between agencies in the non-diagonal elements. As these residuals were identified with the analyst contribution, the matrix can be used to assess the impact of this effect. In particular, it is relevant for this analysis to examine:

- Which agencies incorporate a large analyst contribution into their ratings, and which rely mostly on the public information X;
- The correlation between these analyst effects across agencies, and whether analysts in different agencies behave similarly.

6.2 Computation of synthetic ratings

As indicated in the introduction, one of the purposes of this paper is to propose methods of combining multiple credit assessments. The resulting combination (a synthetic rating) can be useful in several ways:

- It can be used as a benchmark assessment by a supervisor or central bank that does not aim at making its own independent assessment.
- It can provide a solution in cases of split ratings, a need particularly acute when competing ratings are near or under threshold values set by supervisors, central banks or bank counterparties.
- It can help compare new sources of credit assessments to more traditional ones, such as new ECAIs ratings to the ratings of internationally active rating agencies.

Typical combinations of multiple rating assessments could be: the average of the assessments, the maximum (or minimum) of the assessments, the median of the assessments, or the second best (or worse) assessment.[14]

The method used to combine multiple assessments will have a bearing on the credit assessment market. On one hand, using the best assessment could lead institutions to “shop around” for favourable assessments in the market. On the other hand, using a median or “second best” rating, as suggested in Basel II, could force a convergence of assessments, since outlying assessments will be disregarded. A bank will probably have to use a simple and transparent method to produce risk weights depending on all available assessments.

The needs of a supervisor are different and in most cases more complicated than those of a ‘commercial’ user of credit assessments. In evaluating a new source of credit assessment, the supervising authority would need to examine how the assessment compares with benchmarks produced by other, already evaluated or approved sources. However, a sophisticated investor making investment decisions and using several sources of credit information could also profit from the development of a benchmark assessment. Finally, model developers in ECAIs or IRB systems could make use of such techniques. This section aims at proposing how to form such a benchmark, while also suggesting a method on treating split assessments.

The approach taken here is based on the assumption that all available assessments are formed in two steps. First, a score is produced by all assessment systems based on the same publicly and easily available information X. This common information produces a core assessment, which analysts responsible for each assessment system adjust according to additional information, methodological innovation or subjective judgement. Therefore, the assessment combinations in this paper consist of adding the core assessment based on the public information X to a function of the analyst contributions obtained through the decomposition described in Section 5. The function must be selected in a way that reduces the

[14] The standardised approach in Basel II proposes such a combination in the case of multiple assessments.

variance of the analyst contribution and the influence of extreme assessments, while at the same time using the information of all the available assessments.

A simple solution would be to use the average or the median as the above-mentioned function. Alternatively, one can attempt to improve on the given ratings by computing a residual (analyst contribution) with smaller variance as a weighted sum of the regression residuals. Both approaches were applied to the residuals (analyst contributions) of the three rating agencies, resulting in four different methods of combining these residuals:

(a) Mean of the residuals.

(b) Median of the residuals.

(c) Best linear combination of residuals:

Under this approach, the new synthetic residual for bank i will have the form:

$$r_i^s = \sum_{j=1}^{J} c^j \hat{r}_i^j ,$$

where the coefficients c^j satisfy $\sum_{j=1}^{J} (c^j)^2 = 1$.

Equivalently, in vector notation:

$$r_i^s = c'\hat{r}_i , \;\; c'c = 1$$

Then the variance of this synthetic residual is:

$$Var(r_i^s) = c'Sc$$

and the optimal vector of coefficients c will be the solution of the minimisation problem:

$$\min c'Sc \text{ under the restriction } c'c = 1$$

The solution to the above problem for c is the standardised eigenvector corresponding to the smallest eigenvalue of the matrix S (see e.g. Johnson and Wichern (1988), p.64).

(d) Best convex combination of residuals:

Alternatively, to avoid the use of negative weights in the sum, the analysis can be restricted to convex combinations of the residuals satisfying:

$$\sum_{j=1}^{J} c^j = 1 \text{ and } c^j \geq 0$$

and attempt to minimise the variance of the synthetic residual under this restriction. The solution of this optimisation problem can be easily obtained numerically.

Minimising the variance of the analyst contribution appears to be a natural optimisation principle for the following reasons:

1. The individual residuals (analyst contributions) were obtained by least square regressions, i.e. they are already the result of a variance minimisation problem. However, this minimisation was performed separately in each agency case, while the goal here is to combine the different assessments. It is natural to follow the same principle (variance minimisation) to accomplish this goal as well.
2. The problem of combining different assessments is, in essence, one of data reduction: one synthetic assessment has to be chosen to "represent" several ones. This is best accomplished by minimising the average deviation of the synthetic assessment from the other assessments, which results in a variance minimisation principle.
3. By viewing the problem as one of estimation of the unknown "correct" assessment, it is natural to select an estimator minimising variance. In fact, the two simplest ways of combining multiple assessments (the mean and the median) have this property under different distributional assumptions, assuming no correlation in the data.

All four proposals (core + average, core +median, core + best linear combination, core +best convex combination) were applied to the residuals (analyst contributions) of the three rating agencies. The results are shown in Annexes 6 and 7. As expected, the "best linear combination" achieves the minimal variance of the analyst contribution, followed by the "best convex combination". However, the linear combination may give some unusual results. For example, on certain occasions the synthetic rating produced will be outside the range of the three ratings, a result that can not be considered desirable. Annex 7 provides a graphical comparison of the four synthetic assessments to the original assessments of the three ratings for different groups of banks (outliers, top banks in terms of total assets and worst rated banks).

Therefore, the proposal based on the above results would be to use the "best convex combination" to obtain a synthetic rating by adding it to the core assessment based on X. It is also worth noting that the median is found to be a viable alternative to the "best convex combination" since it obtains almost the same variance and is easier to compute.

7. Conclusions

The paper is designed to contribute to the discussion of credit assessments in the light of the renewed interest triggered by the latest Basel II consultative papers. It describes a multiple regression model that attributes the default probabilities corresponding to rating grades to a number of core variables. The observed deviations from this model form the basis of further analysis aiming at (a) exploring factors that caused the deviations, (b) comparing assessments of different agencies, and (c) combining these assessment into a single one which enjoys certain optimality properties.

The potential users of the analysis presented in the paper are quite diverse. Supervisors could use it as a tool to evaluate the quality of credit assessments both under the standardised and the IRB approach, as the proposed method of combining assessments provides them with a way to create benchmark ratings based on multiple sources of information. Credit risk managers can also profit from the methods of the paper both as developers of proprietary credit risk technology and as users of multiple assessments.

The arguments presented lead to the following conclusions:

- There is an advantage in analysing and combining assessments based on default probabilities corresponding to assessment grades. While the three major international agencies do display statistically significant differences in rating a given set of credit institutions, there are no statistically significant differences in the historical default probabilities corresponding to these ratings. Hence, historical default rates naturally "correct" rating inconsistencies.

- There is statistical evidence that agencies' ratings depend on balance sheet information, the country of incorporation and the bank's specialisation. Such information, however, is not enough to duplicate the agencies' methodologies and reproduce the agencies' ratings. The agencies' analysts are seen to deviate from this core rating by fine-tuning it based on additional information they may have, their general rating guidelines and personal assessment.

- The paper suggests that a wealth of information can be derived by decomposing the rating into a core, based on information available to everyone, and an analyst contribution, defined as the subjective interpretation of a bank's creditworthiness, often carried out by analysts and leading to a refinement of a rating. In particular, the impact of the analyst's contribution can be measured and the correlation between such contributions by different agencies can be analysed.

- The extent to which the methodologies used to assess credit ratings and the credit ratings themselves will converge remains an open issue. The analysis presented in this paper, however, not only provides a useful tool in analysing the range of possible ratings, but also measures their convergence. In the context of regulatory capital determination, our approach reduces incentives for "credit rating shopping" by better pinpointing the ratings' range of variation and the causes of their divergence. Splitting ratings into a part based on easily available information and a part linked to the value added of rating agencies helps smooth the variance of credit assessment by multiple sources.

- Finally, and although much work is yet to be done in the area of combining credit assessments, the paper puts forward the idea of using the above mentioned analyst contributions to compute optimal combinations of different credit risk assessments. Such combinations consist of adding to the core assessment based on commonly used information a function of the analyst contributions obtained through the regression decomposition described in Section 5. This function is selected in a way that reduces the variance of the analyst contribution and the influence of extreme assessments while, at the same time, using all the available information.

The need to access and consistently use multiple credit assessments poses a number of difficult problems to the credit manager. In this respect, the consultative papers of Basel II are a welcome addition for two reasons: firstly, because they provide a set of criteria that can be used to evaluate the quality of an assessment, and, secondly, because they trigger new discussions on the challenges of credit assessment methodologies, their analysis, consistency and combination. This paper provides credit risk managers with tools they could use to meet these challenges.

References

Beattie, V. and S. Searle (1992a), Bond ratings and inter-rater agreement, *Journal of International Securities Markets*, summer, p. 167-172

Beattie, V. and S. Searle, (1992b), Credit-rating agencies: The relationship between rater agreement and issuer / rater characteristics, *Journal of International Securities Markets*, winter, p. 371-375

Basel Committee on Banking Supervision (2001a), Consultative Document: The Standardised Approach to Credit Risk, January 2001

Basel Committee on Banking Supervision (2001b), Consultative Document: The Internal Ratings Based Approach, January 2001

Bishop, C.M. (1995), Neural Networks for Pattern Recognition, Oxford University Press, Oxford

Blume, M.E., F. Lim and A.C. MacKinlay (1998), The declining credit quality of US corporate debt: Myth or reality, *Journal of Finance*, Vol. LIII, No 4, August, p. 1389-1413

Brunner A., J.P. Krahnen and M. Weber (2000), Information production in credit relationships in commercial banking, *WP No 2000/10, Center for Financial Studies*, Johann Wolfgang Goethe University

Cantor, R. and F. Packer (1994), The credit rating industry, *Federal Reserve Bank of New York Quarterly Review*, Summer-Fall, p. 1-26

Cantor, R. and F. Packer (1997), Differences of opinion and selection bias in the credit rating industry, *Journal of Banking & Finance*, p. 1395-1417

Carey, M. and M. Hrycay (2001), Parametrizing credit risk models with rating data, *Journal of Banking & Finance*, Vol. 35, p. 197-270

Cisneros, G. G., M. S Garcia, and M. A. Sierra (1997), Determinant factors in the structure of risk groups in financial organizations, WP, University of Barcelona

Crouhy, M., D. Galai and R. Mark (2001), Prototype risk rating system, *Journal of Banking and Finance*, 25, p. 47-95

Ederington, L.H. (1986), Why split ratings occur, *Financial Management*, Spring, p. 37- 52

Estrella, A., S. Park, and S. Peristiani (1999), Capital Ratios and Credit Ratings as Predictors of Bank Failures, WP, Federal Reserve Bank of New York, September 1999

Estrella et al (2000), Credit ratings and complementary sources of credit quality information, Bank for International Settlements, Basel, August 2, 2000

Johnson, R.A. and D.W. Wichern (1988), Applied Multivariate Statistical Analysis, Prentice Hall, Englewood Cliffs.

Moody's Investor Service (2000), Moody's public firm risk model: A hybrid approach to modeling short term default risk.

Moody's Investors Service (2001), Ratios and US Banking Ratings, January 2001

Standard & Poor's (1999), Bank Ratings Criteria, February 1999

Treacy, W.F. and M.S. Carey (1998), Credit risk rating at large US banks, *Federal Reserve Bulletin*, November, p. 897-921

ANNEX 1: SAMPLE OF BANKS

The 67 institutions used in this analysis are listed below in alphabetical order of the database codes. The group comprises 57 banks whose dealings contribute to the derivation of the euro-area inter-bank offered euro rate (EURIBOR). In addition, a number of banks from the US and Japan that are major players at international level were added. The National Bank of Greece was added to represent the 12th member of the European Monetary Union, as at the time of the analysis it was not represented in the EURIBOR composition.

1 ABN Amro Holding NV
2 Allied Irish Banks plc
3 Asahi Bank Ltd
4 Banca Commerciale Italiana SpA, (now Intesa BCI)
5 Banca di Roma
6 Banca Nazionale del Lavoro SpA – BNL
7 Banco Bilbao Vizcaya Argentaria SA
8 Banco Santander Central Hispano
9 Bank Austria AG – IAS
10 Bank of America, National Association
11 Bank of Ireland
12 Bank of Tokyo – Mitsubishi
13 Bank One, National Association
14 Bankgesellschaft Berlin AG
15 Banque CPR
16 Banque et Caisse d'Epargne de l'Etat Luxembourg
17 Barclays Capital Finance Limited
18 Bayerische Hypo-und Vereinsbank
19 Bayerische Landesbank Girozentrale
20 BNP Paribas
21 CDC IXIS Capital Markets
22 Caisse Nationale de Crédit Agricole CNCA
23 Caixa Geral de Depositos
24 Crédit Commercial de France
25 Chase Manhattan Bank
26 Citibank NA
27 Commerzbank AG – IAS
28 Confederación Española de Cajas de Ahorros – CECA
29 Crédit Industriel et Commercial – CIC
30 Crédit Lyonnais
31 Dai-Ichi Kangyo Bank Ltd DKB
32 Danske Bank A/S
33 Deutsche Bank AG – IAS
34 Dexia Banque SA
35 Deutsche Genossenschaftsbank DG BANK - IAS
36 Dresdner Bank AG - IAS
37 Erste Bank der Oesterreichischen Sparkassen AG - IAS
38 First Union National Bank
39 Fleet National Bank
40 Fuji Bank Ltd.
41 Industrial Bank of Japan Ltd
42 ING Bank NV
43 San Paolo IMI
44 KBC Bank NV
45 Landesbank Baden-Wuerttemberg
46 Landesbank Hessen-Thueringen Girozentrale - HELABA
47 Merita Bank Plc
48 HSBC Bank Plc
49 Banca Monte dei Paschi di Siena SpA (Gruppo)
50 Morgan Guaranty Trust Company of New York
51 Natexis Banques Populaires
52 National Bank of Greece SA
53 Norddeutsche Landesbank Girozentrale NORD/LB
54 Norinchukin Bank (The)
55 Rabobank Group
56 Sakura Bank Limited (The)
57 Sanwa Bank Ltd
58 Société Générale
59 Sumitomo Mitsui Banking Corporation
60 SunTrust Banks, Inc.
61 Svenska Handelsbanken
62 Tokai Bank Ltd.
63 US Bank NA
64 SBC Warburg Dillon Reed INc-Warburg Dillon Reed LLC
65 UniCredito Italiano SpA
66 Wells Fargo & Company
67 Westdeutsche Landesbank Girozentrale WestLB

ANNEX 2: HISTORICAL DEFAULT RATES

Table 1: Default rates at 1 and 5 year horizons by agency (percent)

Bond Rating		1 Year Default Rate		5 Year Default Rate	
S&P	*Moody's*	*S&P*	*Moody's*	*S&P*	*Moody's*
AAA	Aaa	0.00	0.00	0.15	0.22
AA	Aal	0.00	0.00	0.27	0.25
AA	Aa2	0.00	0.00	0.11	0.50
AA-	Aa3	0.00	0.07	0.40	0.45
A+	A1	0.03	0.00	0.48	0.75
A	A2	0.04	0.00	0.32	0.66
A-	A3	0.07	0.00	0.82	0.45
BBB+	Baa1	0.20	0.04	1.15	1.45
BBB	Baa2	0.19	0.08	1.36	1.29
BBB-	Baa3	0.30	0.31	3.21	2.79
BB+	Ba1	0.62	0.64	5.79	8.45
BB	Ba2	0.78	0.59	6.88	9.66
BB-	Ba3	1.19	2.55	12.23	20.76
B+	B1	2.42	3.56	16.18	25.56
B	B2	7.93	6.85	24.66	28.52
B-	B3	9.84	12.41	29.16	37.49
CCC	Caal-C	20.39	18.31	41.29	38.30
Investment-Grade		0.08	0.04	0.71	0.82
Speculative-Grade		3.83	3.67	16.08	20.26

Source: Estrella et al, 1999
Note: Moody's data covers the period 1983-1998. S&P data covers the period 1981-1998.

Table 1 above lists the historical default frequencies for S & P and Moody's rated issuers (as percentages) for one- and five-year horizons.

ANNEX 3: CORRELATIONS AMONG THE TRANSFORMED DEFAULT RATES AND THE FINANCIAL RATIOS

Table 2 – Correlation table

	Average Smooth Logit	Tier 1 ratio	Total Capital Ratio	Equity on total assets	Equity on Short-term Funding	Net Interest Margin	Pre-tax Operating Income	ROAA	ROAE	Cost-to-income Ratio	Net Loans on Total Assets	Percentage of Deposits Lent
Average Smooth Logit		-0.4336	0.0774	-0.0343	-0.1217	-0.0752	-0.3569	-0.3337	-0.4503	0.0026	0.3217	-0.1381
Tier 1 ratio	-0.4336		0.5717	0.2876	0.3919	0.1510	0.2796	0.3227	0.3282	0.1373	-0.1457	0.0234
Total Capital Ratio	0.0774	0.5717		0.2072	0.2202	0.0641	0.0290	0.0741	-0.0130	0.1198	-0.0016	-0.0363
Equity on total assets	-0.0343	0.2876	0.2072		0.7358	0.8651	0.6944	0.6617	0.2398	-0.2103	0.4854	-0.0949
Equity on Short-term Funding	-0.1217	0.3919	0.2202	0.7358		0.5906	0.4999	0.4288	0.1135	0.0216	0.1182	-0.0882
Net Interest Margin	-0.0752	0.1510	0.0641	0.8651	0.5906		0.7149	0.6743	0.3159	-0.1493	0.5551	0.0148
Pre-tax Operating Income	-0.3569	0.2796	0.0290	0.6944	0.4999	0.7149		0.9291	0.7212	-0.4485	0.1723	0.0070
ROAA	-0.3337	0.3227	0.0741	0.6617	0.4288	0.6743	0.9291		0.8548	-0.3670	0.1893	-0.0123
ROAE	-0.4503	0.3282	-0.0130	0.2398	0.1135	0.3159	0.7212	0.8548		-0.2369	-0.0547	0.0565
Cost-to-income Ratio	0.0026	0.1373	0.1198	-0.2103	0.0216	-0.1493	-0.4485	-0.3670	-0.2369		-0.0991	0.1217
Net Loans on Total Assets	0.3217	-0.1457	-0.0016	0.4854	0.1182	0.5551	0.1723	0.1893	-0.0547	-0.0991		0.0830
Percentage of Deposits Lent	-0.1381	0.0234	-0.0363	-0.0949	-0.0882	0.0148	0.0070	-0.0123	0.0565	0.1217	0.0830	

Note: Sample size=50

Table 2 above shows the correlations between the transformed default rates (Average Smooth Logit) and the financial ratios initially chosen, based on theoretical considerations, as variables for the principal component analysis. These ratios were chosen as the most likely variables to be correlated with the transformed default rates. The analysis revealed that the information contained in the 11 variables could be summarised and explained by the three components indicated in Section 5.2, namely profitability and efficiency, capital adequacy and liquidity, and loans on total assets.

ANNEX 4: REGRESSION RESULTS

This annex reports on the results of the Multiple Regression conducted on the dependent variables SMOOTH LOGIT (i.e. the transformed default rate of sample banks for the average and for each of the three rating agencies) as defined in Section 4. The relevant dependent variable is indicated in the tables' headers. The following regressors are used: total asset (in natural logarithm), net loans on total assets, equity on total assets, and three dummy variables to capture the effect of a bank being incorporated in either Italy or Japan, or of a bank enjoying a particular type of government guarantee (Landesbank). In addition to the usual statistics, the analysis provides a list of outliers (unusually large residuals).

Multiple Regression Analysis - Dependent variable: Average smooth logit

Multiple Regression Analysis

Parameter	Estimate	Standard Error	T-Statistic	P-Value
CONSTANT	-3.48936	0.848239	-4.11365	0.0001
LOG(total asset)	-0.12371	0.04553	-2.7172	0.0086
Net loans on total asset	0.008572	0.002676	3.20375	0.0022
Equity on total asset	-0.01731	0.005488	-3.15373	0.0025
Du Co JP	0.814815	0.117649	6.92582	0
Du Co IT	0.330579	0.139484	2.37001	0.0211
Du Type Landesbank	-0.63403	0.161804	-3.91852	0.0002

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	9.62417	6	1.60403	16.86	0
Residual	5.61461	59	0.095163		
Total (Corr.)	15.2388	65			

R-squared = 63.1558 percent
R-squared (adjusted for d.f.) = 59.4089 percent
Standard Error of Est. = 0.308485
Mean absolute error = 0.210122
Durbin-Watson statistic = 1.57191

Unusual Residuals

Row	Y	Predicted Y	Residual	Studentized Residual
16	-6.13114	-5.49071	-0.64043	-2.26
21	-6.31088	-5.74691	-0.56397	-2.01
45	-6.31088	-5.62518	-0.6857	-2.35
52	-4.59708	-5.46429	0.867209	3.09
55	-6.31088	-5.53204	-0.77884	-2.7

Multiple Regression - Dependent variable: Fitch smooth logit

Multiple Regression Analysis

Parameter	Estimate	Standard Error	T-Statistic	P-Value
CONSTANT	-2.83696	0.805917	-3.52017	0.0009
LOG(total asset)	-0.15108	0.043227	-3.49495	0.001
Net loans on total asset	0.008239	0.002504	3.2911	0.0018
Equity on total asset	-0.02036	0.005073	-4.01444	0.0002
Du Co JP	0.571722	0.116369	4.913	0
Du Co IT	0.297597	0.126417	2.35409	0.0222
Du Type Landesbank	-0.66488	0.146696	-4.5324	0

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	6.46398	6	1.07733	13.89	0
Residual	4.1897	54	0.077587		
Total (Corr.)	10.6537	60			

R-squared = 60.6737 percent
R-squared (adjusted for d.f.) = 56.3041 percent
Standard Error of Est. = 0.278545
Mean absolute error = 0.187633
Durbin-Watson statistic = 1.61068

Unusual Residuals

Row	Y	Predicted Y	Residual	Studentized Residual
5	-4.41591	-4.99342	0.57751	2.37
21	-6.1086	-5.60106	-0.50754	-2.03
45	-6.1086	-5.52	-0.5886	-2.23
52	-4.65772	-5.32281	0.66509	2.59
55	-6.1086	-5.452	-0.6566	-2.52

Multiple Regression - Dependent variable: S&P smooth logit

Multiple Regression Analysis

Parameter	Estimate	Standard Error	T-Statistic	P-Value
CONSTANT	-3.60252	1.27108	-2.83423	0.0065
LOG(total asset)	-0.13784	0.06697	-2.05823	0.0446
Net loans on total assets	0.01356	0.00375	3.61712	0.0007
Equity on total asset	-0.02139	0.0075	-2.85161	0.0062
Du Co JP	1.08054	0.15515	6.96462	0
Du Co IT	0.36509	0.19842	1.84004	0.0715
Du Type Landesbank	-0.86796	0.2402	-3.61349	0.0007

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	16.5908	6	2.76513	17.32	0
Residual	8.30391	52	0.15969		
Total (Corr.)	24.8947	58			

R-squared = 66.6438 percent
R-squared (adjusted for d.f.) = 62.795 percent
Standard Error of Est. = 0.399613
Mean absolute error = 0.254124
Durbin-Watson statistic = 1.68599

Unusual Residuals

Row	Y	Predicted Y	Residual	Studentized Residual
9	-6.39549	-5.62591	-0.76958	-2.01
30	-4.79573	-5.70463	0.90889	2.41
45	-6.71544	-5.84951	-0.86594	-2.3
52	-4.47578	-5.6735	1.19771	3.43
55	-6.71544	-5.66566	-1.04978	-2.86

Multiple Regression - Dependent variable: Moody’s smooth logit

Multiple Regression Analysis

Parameter	Estimate	Standard Error	T-Statistic	P-Value
CONSTANT	-3.98071	0.81669	-4.87423	0
LOG(total asset)	-0.08621	0.04312	-1.99952	0.0502
Net loans on total assets	0.00422	0.00245	1.72433	0.09
Equity on total asset	-0.0104	0.00496	-2.09734	0.0403
Du Co JP	0.80497	0.10601	7.59324	0
Du Co IT	0.25433	0.1248	2.03797	0.0461
Du Type Landesbank	-0.50242	0.14486	-3.46824	0.001

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	7.17902	6	1.1965	15.73	0
Residual	4.41313	58	0.0760885		
Total (Corr.)	11.5922	64			

R-squared = 61.93 percent
R-squared (adjusted for d.f.) = 57.9917 percent
Standard Error of Est. = 0.275841
Mean absolute error = 0.188847
Durbin-Watson statistic = 1.87827

Unusual Residuals

Row	Y	Predicted Y	Residual	Studentized Residual
21	-6.1086	-5.55891	-0.549696	-2.24
29	-4.89954	-5.47195	0.572413	2.17
35	-4.89954	-5.5005	0.600967	2.29
45	-6.1086	-5.52558	-0.583021	-2.22
52	-4.65772	-5.41	0.752274	3.01
55	-6.1086	-5.49361	-0.614995	-2.36

ANNEX 5: COMPARING RESIDUALS

The tables below present the summary of statistics and correlations of residuals derived from the four regressions detailed in Annex 4

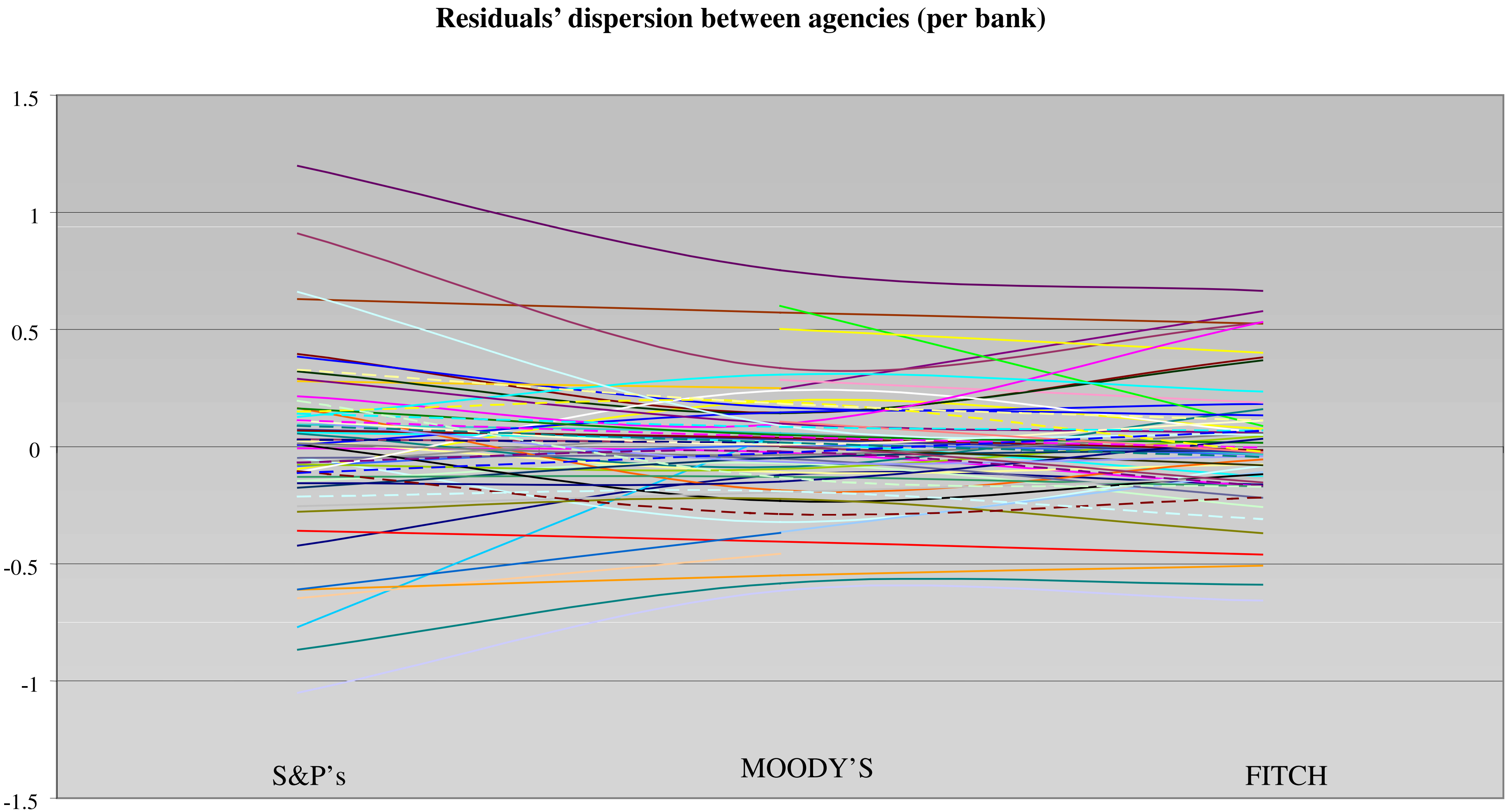
Summary Statistics

	S & P RESIDUALS	**MOODY'S RESIDUALS**	**FITCH RESIDUALS**	**Aver RESIDUALS**
Count	53	53	53	53
Average	0.0313959	-0.0165342	-0.0179457	-0.0100951
Variance	0.130584	0.0593952	0.0693949	0.0761579
Standard deviation	0.361364	0.243711	0.263429	0.275967
Minimum	-1.04978	-0.614995	-0.656603	-0.77884
Maximum	1.19771	0.752274	0.665085	0.867209
Range	2.24749	1.36727	1.32169	1.64605
Stnd. skewness	0.498236	0.443968	0.652259	0.353345
Stnd. kurtosis	4.7883	3.14742	1.54801	3.93802

Correlations

	S & P RESIDUALS	MOODY'S RESIDUALS	FITCH RESIDUALS	Aver RESIDUALS
S & P RESIDUALS		0.8446	0.8480	0.9567
MOODY'S RESIDUALS	0.8446		0.8661	0.9415
FITCH RESIDUALS	0.8480	0.8661		0.9456
Aver RESIDUALS	0.9567	0.9415	0.9456	

Chart 2 – Dispersion of residuals – each line connects the residual relative to the same bank for the three different rating agencies. The chart indicates that the residuals for S&P are more dispersed than for Moody’s and Fitch.



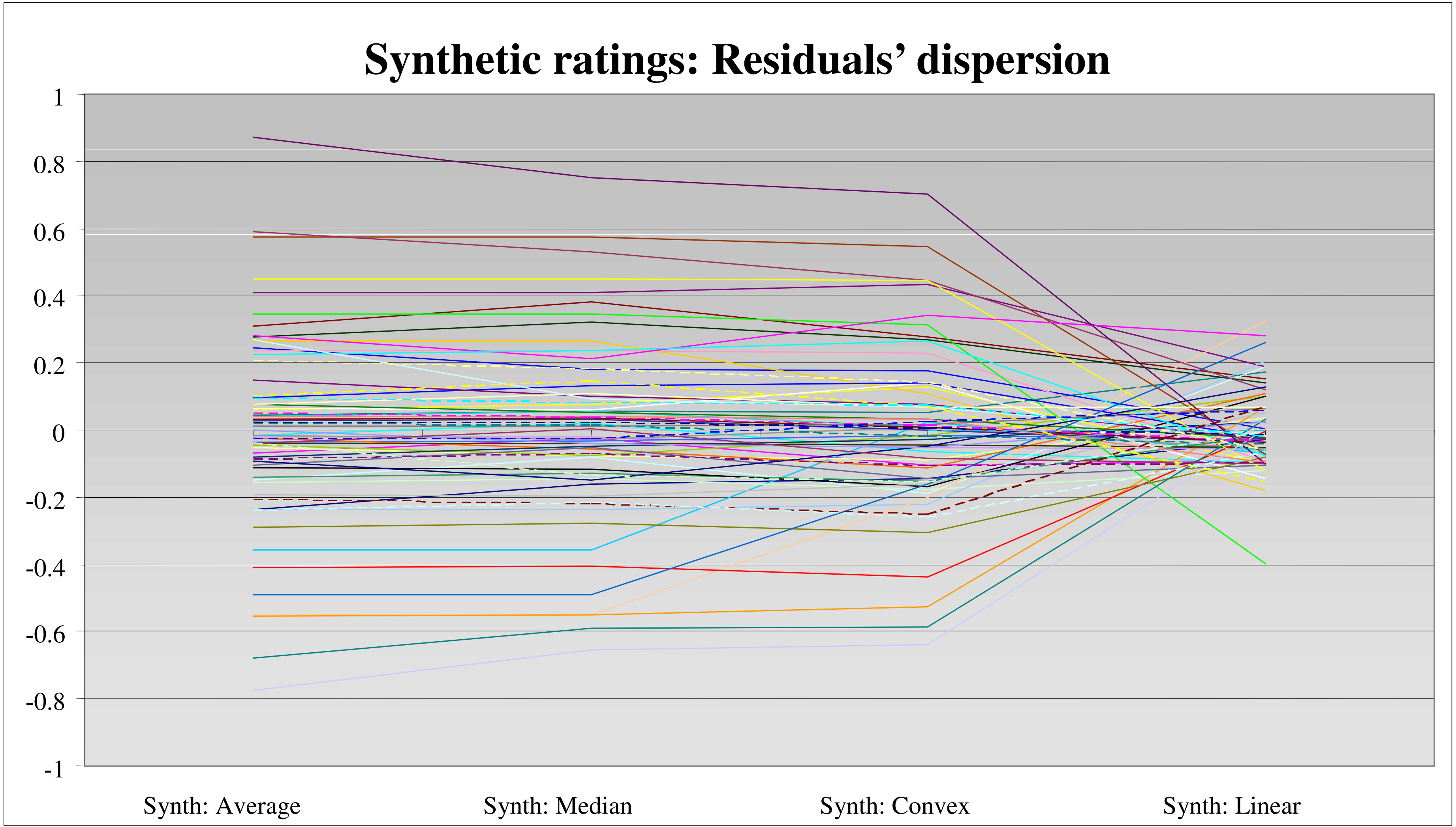
ANNEX 6: COMBINING RATING ASSESSMENTS

Table 3 – Summary statistics for the synthetic residuals

	Synt Average of resid	**Synt Median of resid**	**Synt convex combination**	**Synt linear combination**
Count	66	66	66	66
Average	0.000412708	0.000700900	-0.000000006	0.000000011
Variance	0.081244	0.071428	0.059089	0.014337
Standard deviation	0.285033	0.267260	0.243082	0.119738
Minimum	-0.773793	-0.656603	-0.638379	-0.397889
Maximum	0.871690	0.752274	0.703274	0.326617
Range	1.645480	1.408880	1.341650	0.724505
Stnd. skewness	-0.039478	-0.021501	0.425837	0.804956
Stnd. kurtosis	2.496300	1.622150	2.127990	2.641760

The table above presents the summary statistics of the combinations of residuals proposed in Section 6. A graphical comparison of the dispersion achieved by these four combinations is given in Chart 3. Both the table and the graph show that the optimal linear combination of the residuals achieves the lower dispersion (variance, standard deviation) while the optimal convex combination achieves the second best dispersion.

Chart 3 – Synthetic ratings' dispersion – each line connects the residual relative to the same bank for the four different synthetic ratings



ANNEX 7: COMPARISON OF DEFAULT PROBABILITIES

In the charts to follow, the assessment combinations (synthetic ratings) consist of adding the core assessment based on the public information X to a function of the analyst contributions obtained through the decomposition described in Section 5. The function has been selected in a way that reduces the variance of the analyst contribution and the influence of extreme assessments, while at the same time using the information of all the available assessments. Out of the four possible synthetic ratings proposed, the one based on the best convex combination of the analyst contributions (regression residuals) is selected. The charts show the default probabilities for three sub-groups of banks (outliers, i.e. banks that are rated considerably better or worse than what the values of the regressors would actually predict, top banks by total assets and worst rated banks). The default probabilities shown are those corresponding to the S &P, Moody's, Fitch and synthetic rating.

Chart 4 Default probabilities for outliers: comparison of synthetic ratings

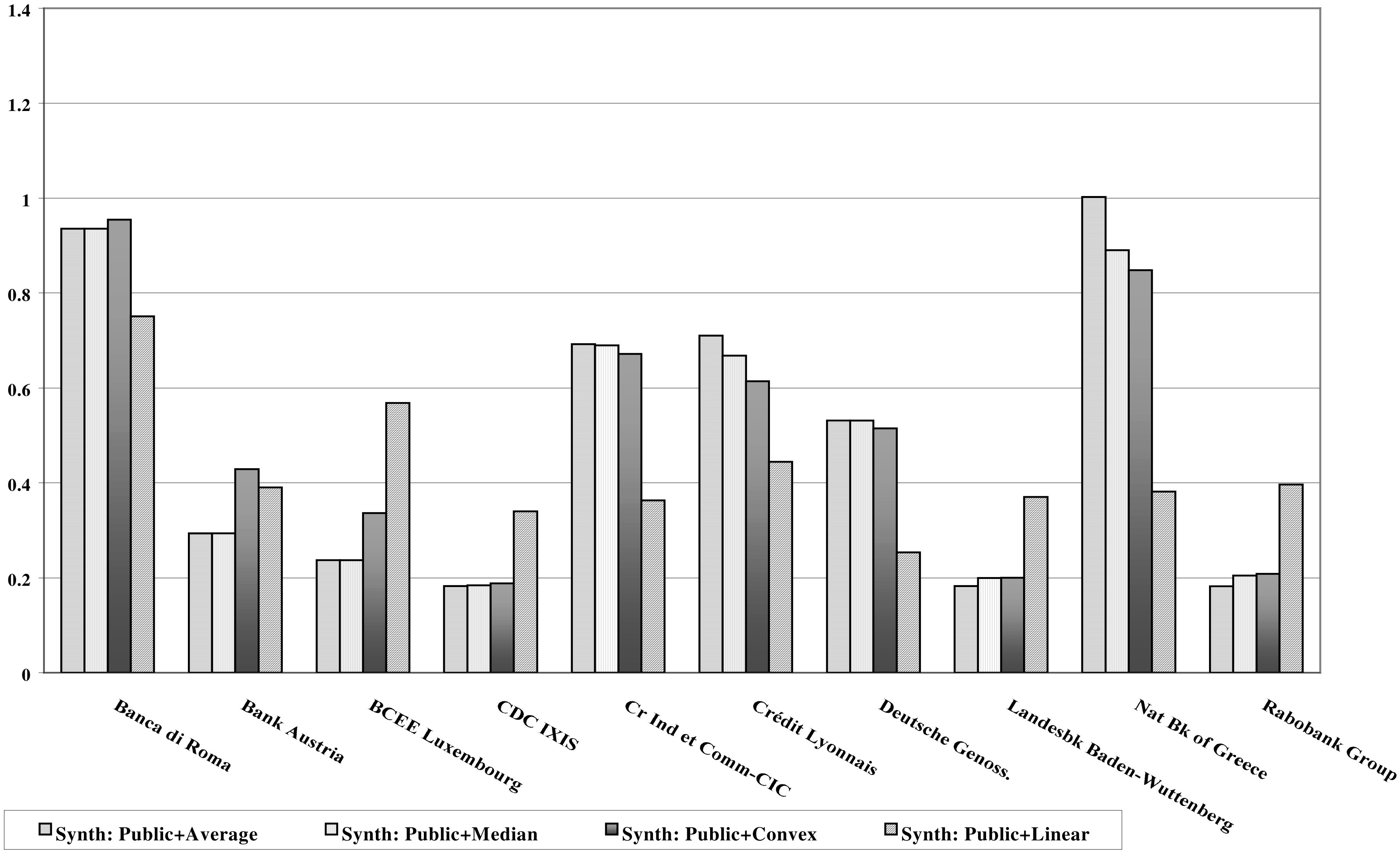


Chart 5 Default probabilities for outliers: agencies versus synthetic ratings

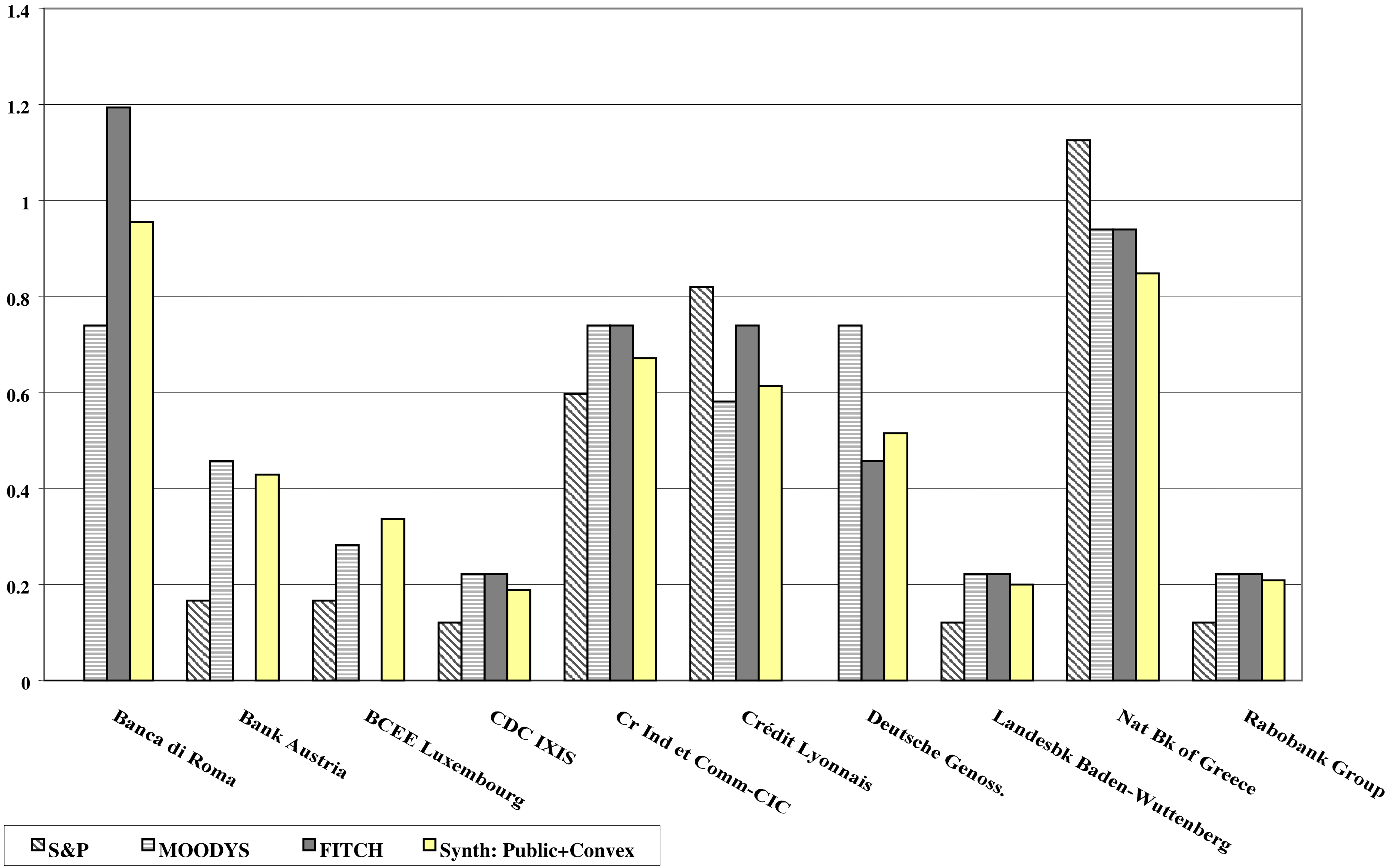


Chart 6 Default probabilities for the Top 10 banks by total assets: agencies versus synthetic ratings

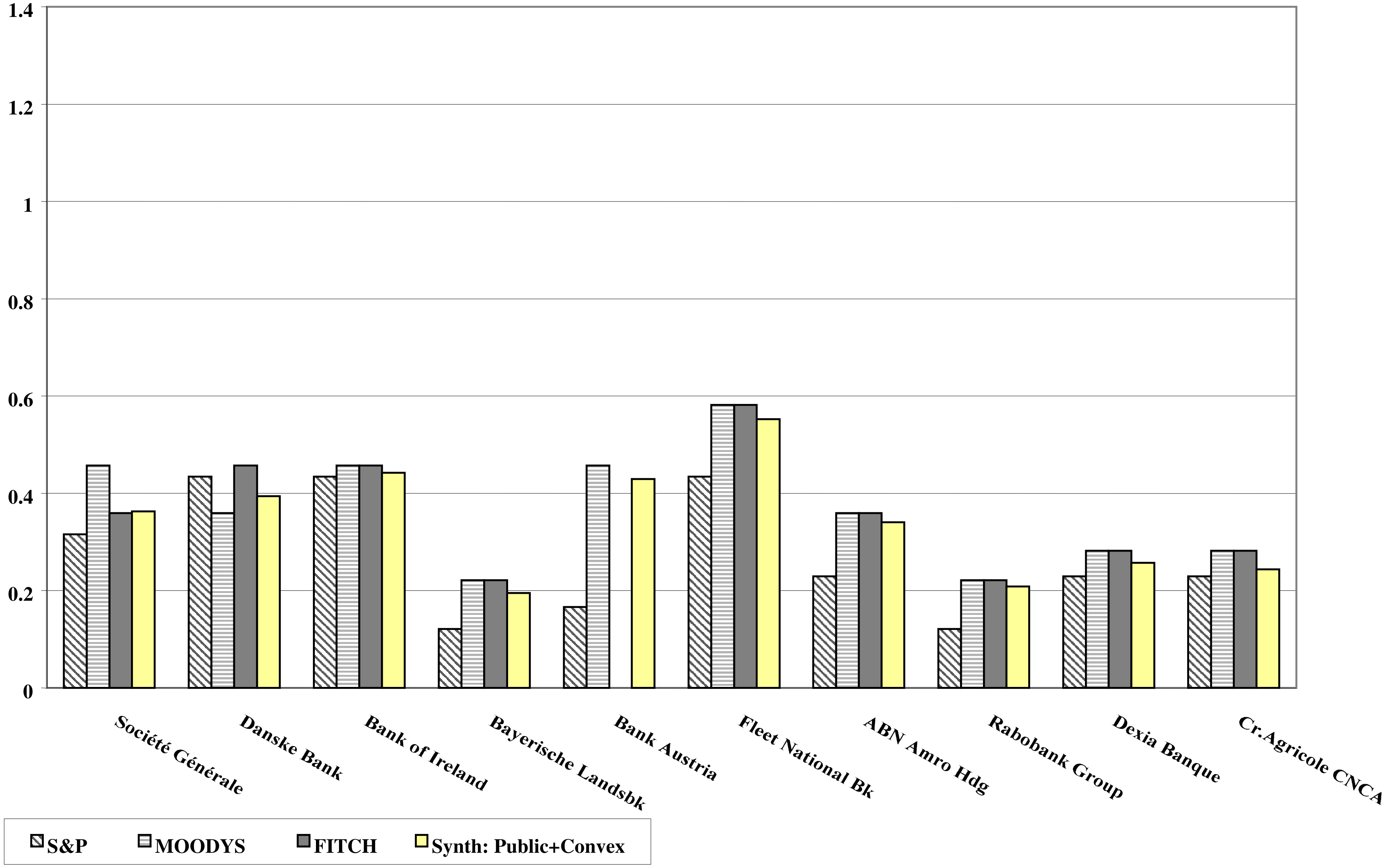
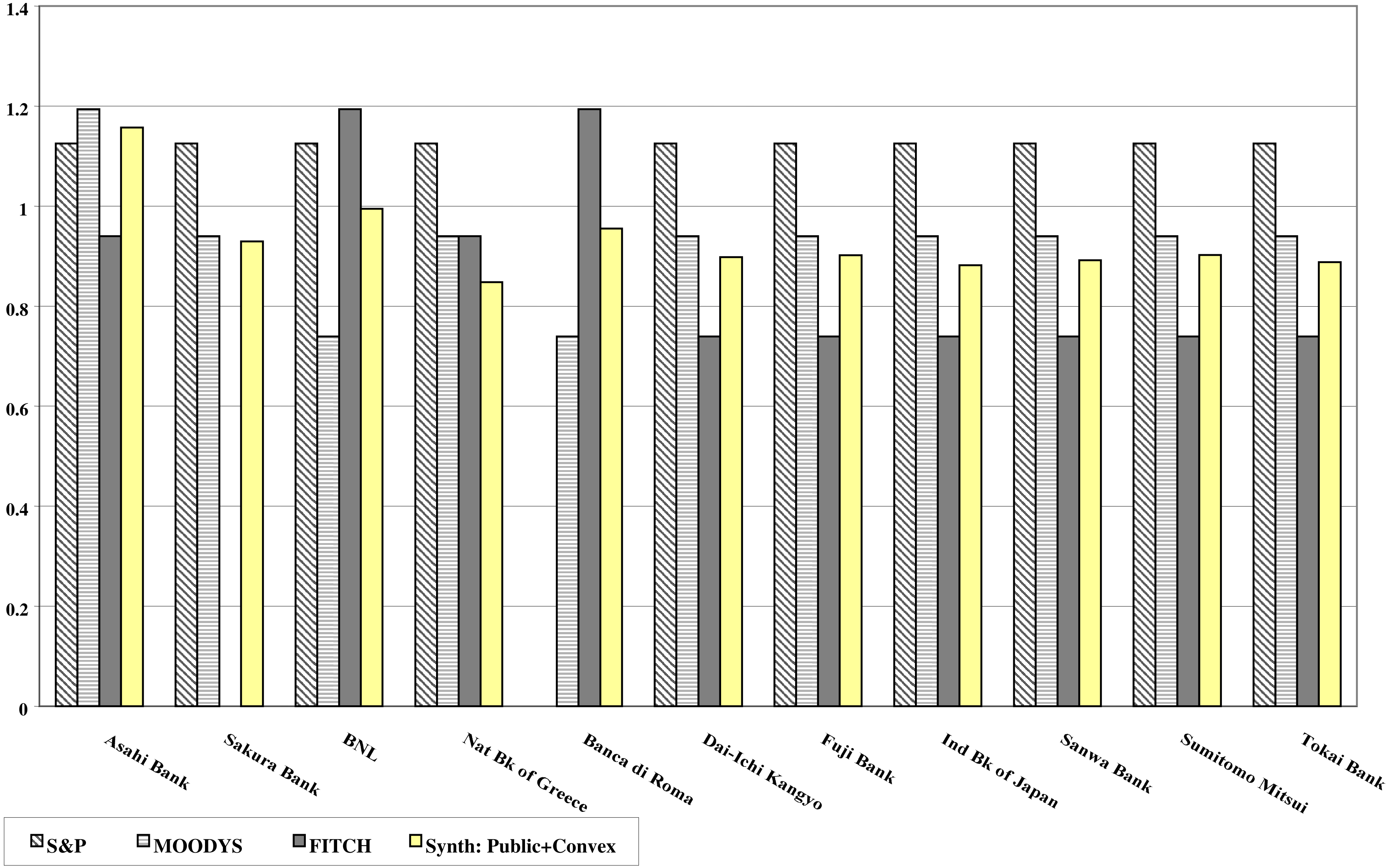


Chart 7 Default probabilities for the 11 worst rated banks: agencies versus synthetic ratings



European Central Bank Working Paper Series

1 "A global hazard index for the world foreign exchange markets" by V. Brousseau and F. Scacciavillani, May 1999.

2 "What does the single monetary policy do? A SVAR benchmark for the European Central Bank" by C. Monticelli and O. Tristani, May 1999.

3 "Fiscal policy effectiveness and neutrality results in a non-Ricardian world" by C. Detken, May 1999.

4 "From the ERM to the euro: new evidence on economic and policy convergence among EU countries" by I. Angeloni and L. Dedola, May 1999.

5 "Core inflation: a review of some conceptual issues" by M. Wynne, May 1999.

6 "The demand for M3 in the euro area" by G. Coenen and J.-L. Vega, September 1999.

7 "A cross-country comparison of market structures in European banking" by O. de Bandt and E. P. Davis, September 1999.

8 "Inflation zone targeting" by A. Orphanides and V. Wieland, October 1999.

9 "Asymptotic confidence bands for the estimated autocovariance and autocorrelation functions of vector autoregressive models" by G. Coenen, January 2000.

10 "On the effectiveness of sterilized foreign exchange intervention" by R. Fatum, February 2000.

11 "Is the yield curve a useful information variable for the Eurosystem?" by J. M. Berk and P. van Bergeijk, February 2000.

12 "Indicator variables for optimal policy" by L. E. O. Svensson and M. Woodford, February 2000.

13 "Monetary policy with uncertain parameters" by U. Söderström, February 2000.

14 "Assessing nominal income rules for monetary policy with model and data uncertainty" by G. D. Rudebusch, February 2000.

15 "The quest for prosperity without inflation" by A. Orphanides, March 2000.

16 "Estimating the implied distribution of the future short term interest rate using the Longstaff-Schwartz model" by P. Hördahl, March 2000.

17 "Alternative measures of the NAIRU in the euro area: estimates and assessment" by S. Fabiani and R. Mestre, March 2000.

18 "House prices and the macroeconomy in Europe: Results from a structural VAR analysis" by M. Iacoviello, April 2000.

19 "The euro and international capital markets" by C. Detken and P. Hartmann, April 2000.

20 "Convergence of fiscal policies in the euro area" by O. De Bandt and F. P. Mongelli, May 2000.

21 "Firm size and monetary policy transmission: evidence from German business survey data" by M. Ehrmann, May 2000.

22 "Regulating access to international large value payment systems" by C. Holthausen and T. Rønde, June 2000.

23 "Escaping Nash inflation" by In-Koo Cho and T. J. Sargent, June 2000.

24 "What horizon for price stability" by F. Smets, July 2000.

25 "Caution and conservatism in the making of monetary policy" by P. Schellekens, July 2000.

26 "Which kind of transparency? On the need for clarity in monetary policy-making" by B. Winkler, August 2000.

27 "This is what the US leading indicators lead" by M. Camacho and G. Perez-Quiros, August 2000.

28 "Learning, uncertainty and central bank activism in an economy with strategic interactions" by M. Ellison and N. Valla, August 2000.

29 "The sources of unemployment fluctuations: an empirical application to the Italian case" by S. Fabiani, A. Locarno, G. Oneto and P. Sestito, September 2000.

30 "A small estimated euro area model with rational expectations and nominal rigidities" by G. Coenen and V. Wieland, September 2000.

31 "The disappearing tax base: Is foreign direct investment eroding corporate income taxes?" by R. Gropp and K. Kostial, September 2000.

32 "Can indeterminacy explain the short-run non-neutrality of money?" by F. De Fiore, September 2000.

33 "The information content of M3 for future inflation" by C. Trecroci and J. L. Vega, October 2000.

34 "Capital market development, corporate governance and the credibility of exchange rate pegs" by O. Castrén and T. Takalo, October 2000.

35 "Systemic risk: A survey" by O. De Bandt and P. Hartmann, November 2000.

36 "Measuring core inflation in the euro area" by C. Morana, November 2000.

37 "Business fixed investment: Evidence of a financial accelerator in Europe" by P. Vermeulen, November 2000.

38 "The optimal inflation tax when taxes are costly to collect" by F. De Fiore, November 2000.

39 "A money demand system for euro area M3" by C. Brand and N. Cassola, November 2000.

40 "Financial structure and the interest rate channel of ECB monetary policy" by B. Mojon, November 2000.

41 "Why adopt transparency? The publication of central bank forecasts" by P. M. Geraats, January 2001.

42 "An area-wide model (AWM) for the euro area" by G. Fagan, J. Henry and R. Mestre, January 2001.

43 "Sources of economic renewal: from the traditional firm to the knowledge firm" by D. R. Palenzuela, February 2001.

44 "The supply and demand for eurosystem deposits – The first 18 months" by U. Bindseil and F. Seitz, February 2001.

45 "Testing the Rank of the Hankel matrix: a statistical approach" by G. Camba-Mendez and G. Kapetanios, February 2001.

46 "A two-factor model of the German term structure of interest rates" by N. Cassola and J. B. Luís, February 2001.

47 "Deposit insurance and moral hazard: does the counterfactual matter?" by R. Gropp and J. Vesala, February 2001.

48 "Financial market integration in Europe: on the effects of EMU on stock markets" by M. Fratzscher, March 2001.

49 "Business cycle and monetary policy analysis in a structural sticky-price model of the euro area" by M. Casares, March 2001.

50 "Employment and productivity growth in service and manufacturing sectors in France, Germany and the US" by T. von Wachter, March 2001.

51 "The functional form of the demand for euro area M1" by L. Stracca, March 2001.

52 "Are the effects of monetary policy in the euro area greater in recessions than in booms?" by G. Peersman and F. Smets, March 2001.

53 "An evaluation of some measures of core inflation for the euro area" by J.-L. Vega and M. A. Wynne, April 2001.

54 "Assessment criteria for output gap estimates" by G. Camba-Méndez and D. R. Palenzuela, April 2001.

55 "Modelling the demand for loans to the private sector in the euro area" by A. Calza, G. Gartner and J. Sousa, April 2001.

56 "Stabilization policy in a two country model and the role of financial frictions" by E. Faia, April 2001.

57 "Model-based indicators of labour market rigidity" by S. Fabiani and D. Rodriguez-Palenzuela, April 2001.

58 "Business cycle asymmetries in stock returns: evidence from higher order moments and conditional densities" by G. Perez-Quiros and A. Timmermann, April 2001.

59 "Uncertain potential output: implications for monetary policy" by M. Ehrmann and F. Smets, April 2001.

60 "A multi-country trend indicator for euro area inflation: computation and properties" by E. Angelini, J. Henry and R. Mestre, April 2001.

61 "Diffusion index-based inflation forecasts for the euro area" by E. Angelini, J. Henry and R. Mestre, April 2001.

62 "Spectral based methods to identify common trends and common cycles" by G. C. Mendez and G. Kapetanios, April 2001.

63 "Does money lead inflation in the euro area?" by S. N. Altimari, May 2001.

64 "Exchange rate volatility and euro area imports" by R. Anderton and F. Skudelny, May 2001.

65 "A system approach for measuring the euro area NAIRU" by S. Fabiani and R. Mestre, May 2001.

66 "Can short-term foreign exchange volatility be predicted by the Global Hazard Index?" by V. Brousseau and F. Scacciavillani, June 2001.

67 "The daily market for funds in Europe: Has something changed with the EMU?" by G. P. Quiros and H. R. Mendizabal, June 2001.

68 "The performance of forecast-based monetary policy rules under model uncertainty" by A. Levin, V. Wieland and J. C.Williams, July 2001.

69 "The ECB monetary policy strategy and the money market" by V. Gaspar, G. Perez-Quiros and J. Sicilia, July 2001.

70 "Central Bank forecasts of liquidity factors: Quality, publication and the control of the overnight rate" by U. Bindseil, July 2001.

71 "Asset market linkages in crisis periods" by P. Hartmann, S. Straetmans and C. G. de Vries, July 2001.

72 "Bank concentration and retail interest rates" by S. Corvoisier and R. Gropp, July 2001.

73 "Interbank lending and monetary policy transmission – evidence for Germany" by M. Ehrmann and A. Worms, July 2001.

74 "Interbank market integration under asymmetric information" by X. Freixas and C. Holthausen, August 2001.

75 "Value at risk models in finance" by S. Manganelli and R. F. Engle, August 2001.

76 "Rating agency actions and the pricing of debt and equity of European banks: What can we infer about private sector monitoring of bank soundness?" by R. Gropp and A. J. Richards, August 2001.

77 "Cyclically adjusted budget balances: An alternative approach" by C. Bouthevillain, P. Cour-Thimann, G. van den Dool, P. Hernández de Cos, G. Langenus, M. Mohr, S. Momigliano and M. Tujula, September 2001.

78 "Investment and monetary policy in the euro area" by B. Mojon, F. Smets and P. Vermeulen, September 2001.

79 "Does liquidity matter? Properties of a synthetic divisia monetary aggregate in the euro area" by L. Stracca, October 2001.

80 "The microstructure of the euro money market" by P. Hartmann, M. Manna and A. Manzanares, October 2001.

81 "What can changes in structural factors tell us about unemployment in Europe?" by J. Morgan and A. Mourougane, October 2001.

82 "Economic forecasting: some lessons from recent research" by D. Hendry and M. Clements, October 2001.

83 "Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts" by K. F. Wallis, November 2001.

84 "Data uncertainty and the role of money as an information variable for monetary policy" by G. Coenen, A. Levin and V. Wieland, November 2001.

85 "Determinants of the euro real effective exchange rate: a BEER/PEER approach" by F. Maeso-Fernandez, C. Osbat and B. Schnatz, November 2001.

86 "Rational expectations and near rational alternatives: how best to form expecations" by M. Beeby, S. G. Hall and S. B. Henry, November 2001.

87 "Credit rationing, output gap and business cycles" by F. Boissay, November 2001.

88 "Why is it so difficult to beat the random walk forecast of exchange rates?" by L. Kilian and M. P. Taylor, November 2001.

89 "Monetary policy and fears of financial instability" by V. Brousseau and C. Detken, November 2001.

90 "Public pensions and growth" by S. Lambrecht, P. Michel and J.-P. Vidal, November 2001.

91 "The monetary transmission mechanism in the euro area: more evidence from VAR analysis" by G. Peersman and F. Smets, December 2001.

92 "A VAR description of the effects of monetary policy in the individual countries of the euro area" by B. Mojon and G. Peersman, December 2001.

93 "The monetary transmission mechanism at the euro-area level: issues and results using structural macroeconomic models" by P. McAdam and J. Morgan, December 2001.

94 "Monetary policy transmission in the euro area: what do aggregate and national structural models tell us?" by P. van Els, A. Locarno, J. Morgan and J.-P. Villetelle, December 2001.

95 "Some stylised facts on the euro area business cycle" by A.-M. Agresti and B. Mojon, December 2001.

96 "The reaction of bank lending to monetary policy measures in Germany" by A. Worms, December 2001.

97 "Asymmetries in bank lending behaviour. Austria during the 1990s" by S. Kaufmann, December 2001.

98 "The credit channel in the Netherlands: evidence from bank balance sheets" by L. De Haan, December 2001.

99 "Is there a bank lending channel of monetary policy in Spain?" by I. Hernando and J. Martínez-Pagés, December 2001.

100 "Transmission of monetary policy shocks in Finland: evidence from bank level data on loans" by J. Topi and J. Vilmunen, December 2001.

101 "Monetary policy and bank lending in France: are there asymmetries?" by C. Loupias, F. Savignac and P. Sevestre, December 2001.

102 "The bank lending channel of monetary policy: identification and estimation using Portuguese micro bank data" by L. Farinha and C. Robalo Marques, December 2001.

103 "Bank-specific characteristics and monetary policy transmission: the case of Italy" by L. Gambacorta, December 2001.

104 "Is there a bank lending channel of monetary policy in Greece? Evidence from bank level data" by S. N. Brissimis, N. C. Kamberoglou and G. T. Simigiannis, December 2001.

105 "Financial systems and the role of banks in monetary policy transmission in the euro area" by M. Ehrmann, L. Gambacorta, J. Martínez-Pagés, P. Sevestre and A. Worms, December 2001.

106 "Investment, the cost of capital, and monetary policy in the nineties in France: a panel data investigation" by J.-B. Chatelain and A. Tiomo, December 2001.

107 "The interest rate and credit channel in Belgium: an investigation with micro-level firm data" by P. Butzen, C. Fuss and P. Vermeulen, December 2001.

108 "Credit channel and investment behaviour in Austria: a micro-econometric approach" by M. Valderrama, December 2001.

109 "Monetary transmission in Germany: new perspectives on financial constraints and investment spending" by U. von Kalckreuth, December 2001.

110 "Does monetary policy have asymmetric effects? A look at the investment decisions of Italian firms" by E. Gaiotti and A. Generale, December 2001.

111 "Monetary transmission: empirical evidence from Luxembourg firm level data" by P. Lünnemann and T. Mathä, December 2001.

112 "Firm investment and monetary transmission in the euro area" by J.-B. Chatelain, A. Generale, I. Hernando, U. von Kalckreuth and P. Vermeulen, December 2001.

113 "Financial frictions and the monetary transmission mechanism: theory, evidence and policy implications" by C. Bean, J. Larsen and K. Nikolov, January 2002.

114 "Monetary transmission in the euro area: where do we stand?" by I. Angeloni, A. Kashyap, B. Mojon, D. Terlizzese, January 2002.

115 "Monetary policy rules, macroeconomic stability and inflation: a view from the trenches" by A. Orphanides, December 2001.

116 "Rent indices for housing in West Germany 1985 to 1998" by J. Hoffmann and C. Kurz., January 2002.

117 "Hedonic house prices without characteristics: the case of new multiunit housing" by O. Bover and P. Velilla, January 2002.

118 "Durable goods, price indexes and quality change: an application to automobile prices in Italy, 1988-1998" by G. M. Tomat, January 2002.

119 "Monetary policy and the stock market in the euro area" by N. Cassola and C. Morana, January 2002.

120 "Learning stability in economics with heterogenous agents" by S. Honkapohja and K. Mitra, January 2002.

121 "Natural rate doubts" by A. Beyer and R. E. A. Farmer, February 2002.

122 "New technologies and productivity growth in the euro area" by F. Vijselaar and R. Albers, February 2002.

123 "Analysing and combining multiple credit assessments of financial institutions" by E. Tabakis and A. Vinci, February 2002.