Philosophy of financial risk models

An article about the current status

by Dr. A. Kukler MFE

December, 2018

Lay out of the article

1. Risk management: 1) awareness, 2) counteraction, 3) evaluation (measurement).
   In this article we work on 1) awareness.
2. Description of current main stream credit risk models ito epistemology & methodology
   Empirical statistics
3. Nozick's description of what this ("probability example") means
4. Negligence of expertise whilst it is fully used in all elements of the chain
5. Underdetermination by data
6. Data quality
7. Are we happy with our models? Positivity metaphor. *Reflective* character.
   What can we do ourselves?

Contents

Introduction

Increasingly, financial risk models are applied in the business environment. Especially banks use risk models to perform main business functions, such as risk management, client selection and deal structuring, capitalisation, performance measurement, etc.

The current mainstream financial risk models are quantitative, based on statistics, and rely heavily, as a condition *sine qua non,* on computing power and communication networks., such as inter- or intranet. This technology is relatively new, and still developing rapidly (Moore's law says a doubling of computing power every two years).

Artificial Intelligence, robotics, deep learning is flowing into bank's offices, pushed by technological developments. The technological asset is developing fast. However, the human factor, currently still the initiator of technological drive, is much slower in its developments. Understanding of the new technology is slower than the technological development itself. Mainstream models are based on statistical theories that were developed in the 17th century[1] or based on rational actor theories that are Modernist (based on Descartes, 1648). In recent years, the most prevalent version of rational choice theory, expected utility theory, has been challenged by the experimental results of behavioral economics. Economists are learning from other fields, such as psychology, and are enriching their theories of choice in order to get a more accurate view of human decision-making. For example, the behavioral economist and experimental psychologist Daniel Kahneman won the Nobel Memorial Prize in Economic Sciences in 2002 for his work in this field. However, proliferation of these new ideas is slow.

In this article, an attempt is made to philosophy about the development and application of financial risk models in order to gain more understanding about the implications of using current mainstream models. This is done in the aftermath of the largest financial crisis in the life of the author, which, not coincidentally happened in the year that financial risk models were used first in banks to calculate required (or regulatory) capital (known as Basel II).

Risk management, creating risk awareness

Leed[2] reports about the epic of Gilgamesh, who voluntarily departed for a journey to face the risk of the unknown in 2800 BC. At these times, a risk taker like Gilgamesh was considered a hero (when he returned). According to Leed, the epic was transcribed in 1900BC, which puts the period in which risk is a recorded concept at four millennia.

Risk is a blend between cognitive statements and normative expressions. Beck[3] shows us that risk not only comprises of cognitive, rational or scientistic statements, but also comprises of normative statements. Science (or knowledge) is required to describe the chains of causes and effects that inflict the risks, but normative and judgemental statements are required to identify the risks, assess risks and evaluate risks. Statements on hazards are never reducible to mere statements of facts. As part of their constitution, they contain both a theoretical and a normative component. Risk statements combine knowledge about chains of causes and effects with normative judgements.

According to Beck, risks experienced presume a normative horizon of lost security and broken trust. Risks are objectified negative images of utopias. Risk determinations are unrecognised, still undeveloped symbioses of the natural and human sciences, of everyday and expert rationality, of interest and fact. Herein lays the essential and momentous consequence: in definitions of risks the sciences' monopoly on rationality is broken. There are always competing and conflicting claims, interests and viewpoints of the various agents of modernity and affected

---

[1] In 1654 Pascal and Fermat discovered the theory of probability, the mathematical heart of the concept of risk.

[2] Leed, E.J., <u>The mind of the traveler, from Gilgamesh to global tourism,</u> Basic books, inc., USA, 1991, p 5,6

[3] Beck, U. <u>Risk society: Towards a new modernity</u>. Sage, London, 1992

groups, which are forced together in defining risks in the sense of cause and effect, instigator and injured party.

Forecasting

Bernstein[4] elaborates the concept of calculated risks in terms of its prerequisites.

(p.121) "[There are] three requisite assumptions – as stated by Jacob Bernouilli- which are critical in determining how successfully we can apply measurement and information to predict the future:

1. full information, in order to know how reliable your sample is;
2. independent trials and uncorrelated risks
3. the relevance of quantitative valuation."

According to Bernstein[5]: "We can assemble big pieces of information and little pieces, but we never get all the pieces together. We never know for sure how good our sample is. We have no full information.

Bernstein holds that there is no rational way to calculate the odds (p.204). Eg. if I interview students about their opinion on society, I can not conclude anything about ***the*** opinion of all people in general, without having all types of assumptions which are not rational because counter the available information concerning only students. We do not have independent trials.

Measurement is a quantification of some real life event. Quantification entails expressing only some dimensions of the real life situation in numbers, not all dimensions. This means that in quantifying the event, information is lost. The importance of the information lost determines whether quantitative valuation is relevant. Examples where quantitative valuation is useless concern matters of happiness, love or sympathy; no unimportant aspects of human life.

Epistemology of current mainstream credit risk models

The shift caused by the broader EC movement is characterised by two new epistemological principles:

I. Financial markets as a price discovery mechanism for credit risks. Similar to other sections of the financial markets, objective price discovery of specific assets is driven by the demand and supply of these assets in perfect markets. Many equally informed participants interact as buyers and sellers and in their joint activity, an objective price for any risk can be established. As this is also the price for which you could actually buy or sell the good, this price is considered superior to any other valuation technique.
II. Empiricism and mathematical logic have been introduced by the statisticians. A real quest for data has emerged. All defaulted credit assets must be utterly registered in the hope to gather sufficient information for back-testing and validation of the credit

[4] Bernstein, P.L., Against the gods: The remarkable story of risk, John Wiley & Sons, New York, 1996.

[5] Idem, p.202. Quote: The information you have is not the information you want. The information you want is not the information you need. The information you need is not the information you can obtain. The information you can obtain costs more than you want to pay.

portfolio models[6]. Senior credit risk managers are confronted with results from regression analysis and multivariate distributed models

The EC movement, specifically where it concerns wholesale credit risk management, relies on various epistemological or methodological disciplines. Again, we will use the Rabobank case to analyse the various epistemological or methodological principles being applied in the EC movement. Undoubtedly, using the Rabobank case will show some features typical for the Rabobank, but will also show a large overlap with approaches used by other banks. However, there is no empirical support for the overlap because in this research other banks have not been consulted.

Within the Rabobank projects we can observe the following set of epistemological principles being used:

1. Statistics or empiricism; based on principles of a loss distribution, on probabilities and consequences in terms of loss amounts, on volatilities, correlations, co-variances and confidence levels, CRPM is fundamentally stated in statistical terms. The heavy weight of statistics also translates in a huge quest for data: classified, categorised registration of credit risk. True means backed by objective data.
2. Econometrics; in order to link the statistical terms to economically interesting questions, econometric models have been built, purchased, and implemented, such as Moodys Riscalc, Credit Metrics or KMV, complemented with internally developed models. These models are true by assumption and logic. They help the advanced credit manager in allocating his risk capital.
3. Financial markets; dependent on price curves, market quotes, liquidity, external ratings, standardisation of finance products and master agreements, structuring skills, calculation and valuation techniques, and last but not least, an actual possibility to trade credit risk, CRPM can be considered as a side-product or off-spring of financial markets. Financial markets are also important for their price discovery function. The market is an epistemological tool, with similarities of common sense knowledge (see below). True means traded.
4. Expert opinion; although the above approaches all rely on the availability of data or standardised information about the risks, this is usually not available in the frequencies needed for proper statistical robustness. Most of the current projects (such as rating or LGD projects) running within Rabobank start with an expert opinion based framework, to be back-tested, feed backed upon by application in practice and improved in due time, when more data and experience becomes available. True means common sense.
5. Tacit Knowledge; Polanyi learns us that the particular knowledge of the real world often remains tacit and that knowledge can have the structure of a skill, requiring lots of personal ambitions, beliefs and experiences. He shows an ontological layering of reality in particulars and comprehensive entities and introduces the principle of marginal control. The body needs to be extended ("indwelling") to include the particulars in order to know the distal. Translating this to credit risk, it means that the particulars we are trying to measure regarding the credit risk in our portfolio can never only by themselves bring into existence the comprehension of the actual credit risk which is an apriori-unknown blend of all the particulars. Only when the particulars form part of the body, and by repetitive interaction

[6] BIS II requires a bank to have available seven years of default data regarding all credit assets in its portfolio. Some banks only started after 2001 to implement BIS II and organise compliance with it. These banks have to retro-actively collect default data, ie. go back in history, in order to be compliant at the planned implementation date of Jan. 2008.

with the subject of knowledge which allows the construction of the skill, it is possible to master credit risk. Credit Committee members appear to master this type of skill and have the required knowledge of counterparties. Truth is the result of a skill.

Statistics / empiricism

Statistics provide interesting and useful algorithms to provide information about groups of entities which are both similar enough to be compared (similar constitutions), and are different enough to make statistical distinctions (different magnitudes). Statistics apply well to similar qualities with different quantities, eg. the length of a soldier.

In order to create some perspectives on the theory of statistics itself, we can distinct various approaches to one of the cornerstone concepts of statistics; the concept of probability. Kyburg and Smokler[7] show that there are distinct sorts of meaning of probability. One sense of the word is chance, or long run frequency, which is empirical, objective and independent of what one knows. It is the real probability. Another sense is actual, refined or justified degree of belief. This is the known probability, or the one that is used in applications of statistics to the real world, it is the one that fuels our models.

Mathematically, probability has a definite meaning; it is simply a non-negative, additive set function, whose maximum value is unity. However, this undefined term in a formal system, this abstract concept doesn't help us in understanding how the notion of probability can be used in e.g. insurance or risk management in banks. This abstract concept must be connected to the real world to become relevant in action. There are essentially three types of connection that have been proposed:

1. the ***empirical***: the empirical or frequentist conception of probability identifies probability with the limit of a relative frequency (there are so many As in B). According to the authors[8]: "The important point is that a probability statement is taken as making an assertion about the world. It may be right or wrong –and it is generally held that we never really know with certainty which it is – but it is a statement, like a statement about lengths or weights, which is either true or false, and for which the evidence is chiefly observational. In order to find out whether or not a probability statement is true, we must make an empirical investigation, and usually this will be a non-terminating investigation of the sort whose results are said (in a non-empirical sense) to be only "probable"."
2. the ***logical***; the logical approach denies that probabilities are empirical statements at all. In extreme, this view holds that probabilities represent logical relations between a proposition and a body of knowledge, between one statement and another statement (or set of statements) representing evidence. Probability statements are as formal as arithmetic statements. Within a given statement and body of evidence, there is only one probability that correctly represents the situation.
3. the ***subjective***; the subjective view holds that probability represents a relation between statements and evidence, but not necessarily a logical one. The value of a probability represents a degree of belief of a person, and, hence is never uniquely defined, according

[7] Kyburg and Smokler, 1964, p. 3-22
[8] Kyburg and Smokler, p. 5

to Kyburg and Smokler. However, they also state[9] that "Of course, in the case in which the evidence logically entails the statement in question or entails its denial, the criteria of ordinary deductive logic are applicable." That must mean that subjective probabilities in their view should only be applied in those cases where it is impossible to do proper counting (to get the right frequencies via empirical investigation) or to apply logical reasoning and provide definite meanings of probabilities.
This subjective theory is a logical theory in the sense that only certain combinations of belief in related propositions are admissible. For example, it is not admissible to wish to lose objects with positive utility under all possible conditions. The person's body of belief should be coherent and consistent. Respectively this means that the distribution of degrees of belief should obey the conventional rules of the calculus of probabilities (e.g. there should be no bet that always loses), and if the evidence entails the statement the person should have the highest degree of belief in that statement.

According to Oldenburg[10] "Most of the important theoretical results in financial economics which involve preferences are based on expected utility theory, which presupposes that individuals make rational choices according to a time separated utility function."

This means that the theory assumes specific behaviour, how people *should* act. The theory is predominantly of a normative character. Behavioural finance is a new area of research which tries to provide theories that more accurately describe how people *do* act.

Daniel Kahneman and Amos Tversky in 1979-1981 have proven in empirical experiments that human behaviour may be subject to structural deviations of the rational model. According to Oldenburg[11]: "Among the most important fallacies which have been reported are that people treat losses and gains differently and that they are highly sensitive to the format in which a problem is presented to them." People tend to evaluate gains and losses with respect to some reference point, rather than evaluate based on final assets. It seems that people react more to changes in wealth than to expectations of final wealth. A loss is treated differently when previously a gain was made compared to a previous loss. A certain outcome may imply abject poverty for one person or great riches for another, depending on current assets. Also, apparently, individuals are limited in their ability to comprehend and evaluate extreme probabilities, which results in either ignoring or overweighting of highly unlikely events. Next to that, people tend to be loss averse, instead of purely risk averse. A loss (negative change of wealth) represents more value/utility than a profit (positive change in wealth). Finally, people are sensitive to the evaluation frequency in their decisions. That is, an investor might be more concerned with the period to the next moment of evaluation, than with his real investment horizon. For example, pension fund investors are evaluated quarterly and have to show good results per quarter, while their real horizon may be more than 30 years.

These empirical findings conflict with the basic assumptions of expected utility theory, such as:

- Cancellation: this principle states that any state of the world that yields the same outcome regardless of the actual choice that will be made, may be cancelled or eliminated; the choice of states depends only on those states where these prospects yield different outcomes.
- Transitivity: this assumption is needed to represent preferences on an ordinal scale. If A is better than B, and B is better than C, then also A must be better than C.

---

[9] Kyburg and Smokler, p. 17
[10] Oldenkamp, 1999, p. 33
[11] Oldenkamp, p. 34

- Dominance: this assumption states that if a choice is better in at least one state of the world and at least as good in all other states, then that choice should be preferred over all others.
- Invariance: this assumption states that different descriptions of the same set of possible prospects yields the same choice.

Actual behaviour of people shows a deviation from each of these assumptions required for conventional statistical models. The assumptions cause a bias between model results and actual behaviour.

In mainstream statistics, in order to provide interesting, reliable and accurate results, statisticians require data, samples, population demarcations, etc. Statistical knowledge, therefore, is dependent knowledge, conditional knowledge. Statistical conclusions depend on the availability (in terms of quantity, quality and relevance) of data.

Nozick[12] elaborates on this also: "The probability of a statement or of an event [eg. Default event] provides an example of something that is relative. The probability of a statement is relative to evidence. That probability will vary with different evidence, and that probability is not detachable from the evidence as something that holds as a freestanding fact. And the probability of an event having a certain property is relative to a reference class. Different classes into which that event falls will show differing percentages of events having the property in question. To speak of the probability (period) of a statement or of an event, we have to take as given or to hold constant the evidence or the reference class. Indeed, this is not enough. Rather, we must speak explicitly of the probability of a statement relative to given evidence or the probability of an event relative to its being a certain type."

With this in mind, we should consider the many different individuals involved in the credit process, the many different registration methods, the division of work between credit analysts, who "know" how the counterparty is doing, and credit control staff who keep the records and files tidy and up to date. Latter officials are usually the ones who have to provide the basic data for the statistical modelling to a central unit. This central unit consists of econometricians and mathematicians. They transform the data into intelligible graphs and reports for senior management who takes this information into account in making strategic and tactical decisions. For the final users of the statistical information, all Nozick's conditions will be out of horizon, sublimated in the formal end report.

Cools[13] explains about the rhetoric of economics when he states that facts do not speak for themselves. "Statistics, for example, appear as hard numbers but are artefacts contingent upon theory, concept formation, collection technique, and statistical processing techniques. That only statistically significant results get published has long been a scandal among statistical purists: they fear, for example, with some reason that at the five percent level of significance something like five percent of the computer runs will be successful. Moreover, statistical significance seems to give a standard measurement by which to judge whether a hypothesis is true or false, that is independent of any tiresome consideration of how true a hypothesis must be to be economically true enough…The standard used is the irrelevant one of statistical significance…There is no

[12] Nozick, p.17
[13] Cools, p. 28

‘absolute sense’ in which a description is good or bad. The sense must be comparative to a standard, and the standard must be argued economically. Significance in statistics, however useful as an input into economic significance, is not the same thing as economic significance.[14]”

An example within CRPM theory of the above criticism of Cools involves the use of normal distributions to model credit portfolio losses. Normal distributions work quite well around the mean, but not in the far tail of the distribution. However, capital is held for losses in the far tail of the distribution. It thus seems that the normal distribution, however easy to calculate with, counteracts the purpose of the calculations.[15]

Although convenient for statisticians, the application of normal distributions should disturb their message, but this is hidden by the rhetoric power of the statistical language. In the EC framework we are currently implementing, the public assumption is that we reserve enough capital for a 1 in 10.000 years event. On first instance, this sounds very safe. But it could mean that we are hit by extreme losses three years in a row, which could kill the bank. The only thing the bank can subsequently retrieve from the statisticians is the remark that these three years would only appear once every zillion year according to their normality assumptions, but unluckily enough that happened right now.

The fourth remark regarding statistics stems from an analogy with Quantum mechanics. Nozick[16] shows that the discovery of Quantum Mechanics (QM) makes it very plausible that truth is relative:

“Quantum mechanics has led us to maintain that truth is relative to a time. And the considerations that led to this conclusion, when consistently pursued, lead to the further view that truth is relative to a time and place. Truth is relative to spatiotemporal position. Spatiotemporal position is a surprising and unexpected factor in the context of truth, and all

---

[14] Mc Closkey (1985, p202) notes:” The appeal is part of the rhetoric of statistics. The British inventors of statistics, as recipients of classical educations, were skillful in naming their ideas. As William Kruskal, a statistician of note, has argued: ”Suppose that Sir R.A. Fischer – a master of public relations- had not taken over from ordinary English such evocative words as ‘significant’, ‘efficient’, and ‘consistent’ and made them into precisely defined terms of statistical theory. He might, after all, have used utterly dull terms for those properties of estimators, calling them characteristics A, B and C …Would this work have had the same smashing influence that id did? I think not, or at least not as rapidly.”

15 From an RI internal memo dd. Dec 10 2002: “For the modelling of correlations, OWC (the bank’s consultant for CRPM) assumes a multivariate normal distribution for the ease of calculation. The problem with this approach is that it rests on two assumptions which have empirically proven wrong:

1. The normal distribution underestimates the probability of default of each borrower (the tails are too thin).
2. The multivariate normal distribution assumes independence of defaults for each borrower (because the correlation between extreme shocks to asset values disappears in case of the Normal dependency). Especially in cases of collective debtor events, affecting more than one borrower, borrowers behave collectively, ie. change in credit quality collectively or correlated, instead of independent.”

[16] Nozick, p. 43

spatiotemporal positions are equally good. So the present view counts as *relativism about truth*. It might be called the *Copenhagen Interpretation of truth*.[17]"

QM brings forward this relativism because of several effects discovered in QM, such as the:

- Influence of measurement; because we are measuring, we are playing a role in the phenomena to be measured. On the level of particles, this may lead to a change in the behaviour of the particles.
- Disappearance of evidence; ao. due to the role of measurement in a process, but also due to peculiarities of processes, some states of the process might be changed after first observations in a irreversible way, destroying the evidence.

Both effects may sound relevant only on the level of the smallest particles where QM is valid. However, as higher level phenomena consist of particles, laws of QM may well apply, at least to all physical facts according to Nozick[18].

Next to that, also sociologists / anthropologists have accepted the role of the participating observer already a long time ago. In terms of the PoT, the traveller himself is part of the environment, has both an active and passive presence.

Furthermore, it is clear that the bank does play a role itself in generating credit losses, even if the bank does not select any new credits, ie. considering the existing portfolio. In general, it is believed within the bank that early involvement of the bank in case of credit quality deterioration will help the bank in reducing its losses.

Finally, a large problem of all banks implementing CRPM models is the accurate calibration of parameters. Although all evidence was available once to the bank, a lot of information required for calibration was not properly recorded, leading to all sorts of shortcuts, workarounds and other less accurate solutions for the calibration problems. Similar to the levels of particles in QM studies, also in credit risk management evidence tends to get lost, tends to disappear over time, ao. due to registration procedures, such as clearing files of old data.

For any specific asset, one can assume that at least one employee once knew exactly what happened, ie. that full information has been available to the bank. For example, the account manager once knew what he has discussed with the client, the credit risk analyst once knew the right financial figures of the client, as well as its estimated recovery value and the eventual final amount lost on the position. However, this information is lost later because it was not properly recorded according to a central standard and maintained. Evidence, once available to the bank, has disappeared.

Use of statistics for specific risk

It can be questioned whether concepts like average and standard deviation are applicable to specific risk. Taking one extreme of portfolio composition, it becomes clear that the statistical concepts have nothing to offer here.

Eg. when I have a portfolio with only a black customer and a white customer, both can have good track records, while both deviate significant from the average attribute of good customers, ie.

[17] After the Copenhagen based physicist Niels Bohr who holds a similar interpretation of QM.
[18] Nozick, p. 35

grey. The average of the portfolio – grey – does not exist, while both existing customers will have a very large standard deviation, ie. are looked upon as very risky, while they have been showing good track records up till now.

Furthermore, as specific risk entails the risk which is specific to a particular counterparty (also called idiosyncratic risk) it can be questioned whether this can be compared to the idiosyncratic risk of other counterparties to derive statistical measures of risk. For example, an internal investigation into the predictability of the rating of counterparties[19] for defaults revealed that by far most defaults were caused by fraud. This was never factored in the rating and, hence, is not measured. Furthermore, if fraud is so prevalent, one could question the financial numbers that were used to make the required estimations of solvency, liquidity and profitability, which are required to derive a rating.

As stated above, statistics are useful for groups of entities which are both similar enough to be compared (similar constitutions), and are different enough to make statistical distinctions (different magnitudes). To apply them to debtors assumes that debtors are all similar enough. This is not a rational assumption, because we don't have the data to back that up, and traditionally we know that every debtor was treated unique (the tailored suit approach).

Given the enormous variance in conditions and characteristics of credit risk positions in wholesale banking finance (translating in many differences in magnitudes), the amount of required but missing data and the importance of the actual behaviour of people in business context applications of statistical theory (such as EC), the subjective approach to probabilities seems most adequate. If we can not simply count the required evidence or if we can not definitely determine logical relationships, then we have to resort to degrees of belief and accept the fact that people do not behave as they should, do not obey the rational model, but show loss aversion and misconceptions of probability.

Given the ontological features of wholesale finance credit risk, as discussed in § 8.2.3, especially the lack of similarity and large differences in magnitudes, I have serious doubts about the quality of the connection of statistics to the real world, especially when the empirical connection is prescribed, as is in mainstream thinking today. It is evident that the Law of Large Numbers does not apply in wholesale finance. Improvements with respect to the quality of this connection ("the calibrations") will increase the relevance for actions of these models considerably.

Econometrics

Econometrics can be defined as the statistical application of economy. Econometrists built statistical models for economic issues. In fact, econometrics is focused on the connection of mathematics with the real world. In order to analyse econometrics, we will focus on the Rabobank Group econometric model for EC and benefit from an internal analysis.

---

[19] The Financial Risk Score, FRS, as used formerly in RI and DLL to be precise. The FRS provides a rating for the creditworthiness of the (corporate) counterparty.

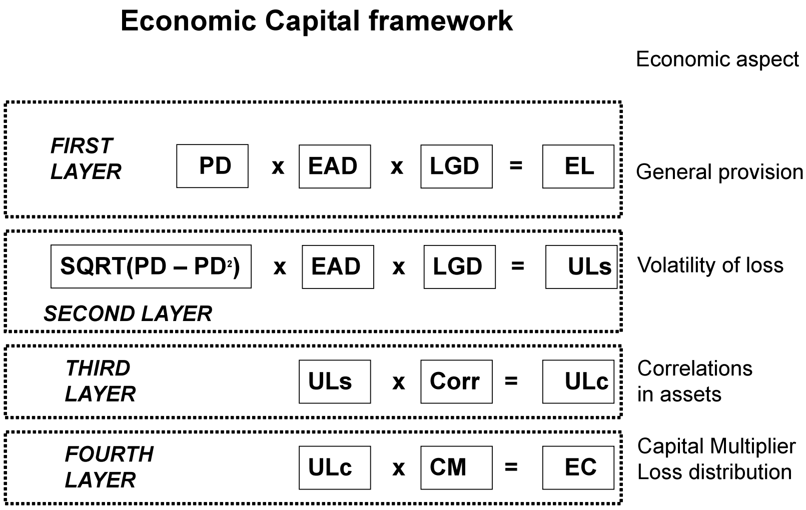
During 2002, the contours of the Rabobank Group policies and standards have become clear in the sense that a Group Credit Risk Methodological document has been issued, to which Business Units (BUs) must adhere while building their specific frameworks. This Group model is also used to calculate Economic Capital and Raroc on a BU level.

Modelling & Research (M&R, a risk management department within Rabo) has analysed the credit risk framework for Rabobank Group as proposed by Oliver Wyman & Co (the consultant) in its paper of February 2001. M&R consists of employees experienced in econometric modelling of financial risks, especially market risks, and employees experienced in credit risk portfolio management from the former portfolio management desk within the former International Credit Department. This analysis will provide a birds-eye overview of the most important comments on the current Credit Risk Model for Rabobank Group

Focus

For this analysis, the EC framework was divided into four layers, which represent different moments of the credit loss distribution.

See the picture below for a definition of the layers:



The analysis is focused on three aspects of the implementation of the EC framework:

1. The first layer of risk factors in the model. The first layer consists of elementary credit risk measurement concepts, such as ratings and probabilities of default, estimations of future exposure amounts (EAD), and estimations of the potential losses in case of default (LGD). The product of these three concepts is called expected loss, and indicates the most likely loss of a portfolio of credit exposures. The first layer already receives a lot of attention in various rating improvement projects and several EAD/LGD modelling

projects. Flaws in the current model include the restriction to a one year horizon and corresponding negligence of long term risk profiles and time value. Furthermore, the current modelling of derivatives exposure raises some critical comments in terms of unexpected exposure and correlations between risk components. Latter has become more important because recently it has been determined that also the counterparty risk in the trading books will be measured by an internal rating based approach and probably also the credit risk in the Credit trading books.

2. The second, third and fourth layer of risk factors in the model. These layers consist of concepts of volatility of losses, correlations between exposures and capital multipliers, of which the product is the economic capital required for the portfolio. Actually all components of the second layer do not measure what they should measure. Either they are currently modelled as benchmarks, provided by the consultant, or they are modelled in a very limited way.
3. The third aspect concerns the embedding of the EC framework in the organisation, ie. the implementation of EC in day to day credit risk management. Comments concern back-testing of parameters, definition of credit risk appetite (portfolio limit setting) and allocation of limits.

Risk management consists of risk control, risk allocation and risk evaluation.

Items 1 and 2 above refer to risk control, which is the measurement and monitoring of risk. Item 3 refers to risk allocation (the allocation of risk limits) and risk evaluation (the assessment of the risk return trade off). See appendix for a full analysis of the econometric model.

Conclusion: it is clear that the connection to the real world created in Rabobank's framework is far from ideal. Many important parameters in the model are not sensitive to the risk components they are supposed to measure. This does not deliver a risk sensitive measure of the credit risk of the portfolio, especially where it concerns the portfolio aspects.

However, first feedback from implementations of the current version of the framework indicate that already this version of the portfolio model is a big step forward and certainly helps in structuring and improvement of credit risk measurement, registration and reporting. The implemented model is a relative improvement but not absolutely reliable, and does not deliver what it promises. However, further versions will be improved by the steep learning curve that the bank is undergoing.

Negligence and widespread application of expert knowledge

Underdetermination by data

Final thesis p38: Theories are underdetermined by the data.[20]

Regarding the underdetermination of theories by the data, Nozick[21] explains that "our observational data are one small consequence of the laws that hold, and sometimes are quite distant from the most basic processes. For things to be different, our observations and data would have to be "in the round" and deep. In that case, although we could not get very far beyond our observations, we would not need to, for those observations would reach all the way to the basic structure of the world; they would be observations *of* the basic structure of the world, a direct experience of the underlying laws, of elementary particles, and of the structure of space and time. In that case, science would not exist – it would be unnecessary. We would know its results already, *by observation*."

Final thesis p231: Herbert Simon[22] has criticised the idea of this rationality [algorithmic rationality[23]]. According to him rationality is bounded, which means that we only calculate scenarios which are known to us according to rules which are disputable. But Schipper[24] mentions two more interesting comments:

1. Scientific theories are empirically underdetermined. Recall that this issue was also mentioned as one of the complicating factors of modernist science in § 5.2.3. An empirical test of a theory involves much more than a simple verification or falsification of the theory. As Nozick pointed out, an empirical test also requires theories for example about the propagation of light, or the working of our sensory organs, etc.
2. Knowledge is algorithmically underdetermined. There are no imperative rules for evaluation of theories. Again, Nozick pointed at this issue in § 5.2.3., when he stated that ideal theories have conflicting properties, such as scope, simplicity, accuracy, etc. Whenever an algorithm is applied, already two moments of judgement must have been passed. First of all in the construction of the algorithm; in defining the risk to be captured, defining its components, defining the way we can measure it, underlying definitions, etc. Second, in the assessment whether a concrete situation fits within the scope of the algorithm. For example, in the EC framework within Rabobank, analysts must first identify the proper rating model to assess the PD of the client, next, a different choice must be made for LGD models. For all these models, an analyst may do an *override* of model generated output if the analyst judges it inapplicable in this case. There may be more than one model applicable if an actual client operates in the grey areas on the demarcation lines between the models.

---

[20] That is, more than one theory can explain the data. (p. 111…), Nozick, R., Invariances, the structure of the objective world, Harvard University Press, 2001, Cambridge, Massachusetts.

[21] Nozick, p. 112

[22] Simon, H.A., The new science of management decision, Harper & Row, New York, 1960

[23] Decision theory is an example, with its theory of expected utility. In this theory an actor is assumed to have a well defined utility function at his disposal which enables him to assign value to all possible scenarios. On the basis of full knowledge of all options to act, of causality structures and probability density functions, the optimal action can be calculated.

[24] Schipper, F., Zin in organisatie, Boom Amsterdam, 1993

Therefore, Schipper also distinguishes the judgemental rationality. Following the English philosopher Brown[25], a **judgement** is then defined as:

> "the ability to evaluate a situation, assess evidence, and come to a reasonable decision without following rules".

Note that judgemental rationality is negatively defined, as being still rational although no algorithms will be applied. The question is then how the reasonability of a decision is to be assessed. Schipper suggests that more general prescriptions would still apply in this rationality. Examples include the pursuit of consistency in judgement, or the requirement to take into account all relevant information.

Next to that, this rationality is linked to the intention to justify the judgement, possibly even by referring to algorithms.

For CRPM, the subjective assessment of creditworthiness of debtors must be expressed in a rating which indicates the chance of defaulting. In order to do so, all aspects which are relevant for the creditworthiness of the debtor must be summarised in one number which can subsequently be compared with the rating of the other debtors.

One could question whether the rating is an under-determination of creditworthiness in the sense that one figure cannot represent the wide array of possibilities to default. For a very simple example, two counterparties can, at a given time, have exactly the same rating, but differ in credit risk widely in case one counterparty is improving its creditworthiness and the other is deteriorating. Latter was good, but is going down, while the first was bad but is improving. At a certain moment, these two will have the same rating, whereas it is clear that the deteriorating counterparty incurs more risk. Recalling the Heisenberg principle, in one number one can not measure position and speed simultaneously. One number is easy for calculations, but may be too easy for risk management.

[25] H.F. Brown, Rationality, London/New York, 1990.