A Darwinian View on Internal Models

Paul Embrechts (1)

(1) **RiskLab and Swiss Finance Institute, Department of Mathematics, ETH Zurich, Raemistrasse 101, 8092 Zurich, Switzerland**

ABSTRACT

Current discussions on regulation within banking (Basel III and IV) as well as insurance (Solvency II and SST) are partly dominated by the way in which internal models may be used for the calculation of regulatory capital and solvency ratios. In this paper, I review these discussions from a historical, personal and an academic point of view. In particular, I pay attention to the internal versus standard model debate taking account of the various technological changes currently facing the financial industry worldwide.

Keywords Basel Committee on Banking Supervision, Charles Darwin, EBA, Expected Shortfall, Extreme Value Theory, Incisive Media, internal models, machine learning, model risk, ORSA, Quantitative Risk Management, RiskLab, Risk Waters Group, robustness, Solvency II, standard models, Value-at-Risk.

About the title

Throughout the paper, I use the Darwinian notions as metaphors like “survival-of-the-fittest-model”, or indeed “the-model-that-best-adapts to market change”, that played through my mind when, many years ago, I started reading the various Pillar I guidelines underlying banking and insurance regulation. Indeed, within the universe of internal models, industry was always looking for the best model. Having said that, I am surely not pushing for a deeper comparison, nor should the allusion to Darwinian concepts distract from the main themes of my discourse.

Introduction: the early days

First of all, it is a great pleasure for me to be able to contribute to the 20th anniversary celebration of the Journal of Risk (JoR) which prompts me to start by

detailing my own publication-footprint related to JoR. No doubt, the year 1998 was a particularly memorable one, the downfall of Long-Term Capital Management (LTCM) raised serious concerns about the financial industry's capabilities in controlling the risk embedded in complicated arbitrage models. I still vividly recall learning about the details of LTCM's rescue while participating in a risk related conference at the University of Cambridge. Around the same time, we published a paper [1] in Risk Magazine stressing the need for industry to start looking more seriously "beyond the bell-curve", i.e. considering non-Gaussian models which allow for a more realistic view on heavy-tailedness to assess risk in financial markets, and hence are better suited from a Quantitative Risk Management (QRM) point of view. Already in 1997 we published a book [2] with the telling title "Modelling Extremal Events for Insurance and Finance" in which we provided the mathematical theory for extreme events - Extreme Value Theory, EVT - needed for looking beyond the Gaussian horizon. On June 5, 1997, I gave a talk on the topic of [2] at the RISK '97 Annual Conference in Chicago. I still fondly remember that at that meeting I held the first physical copy of [2] in my hands and proudly showed it to the audience. Risk Waters Group shortly after asked me to edit a volume [3] within their RISK Books series fully devoted to the topic of extremes in Integrated Risk Management. Browsing through its chapters today, I feel proud about the visionary content provided by many excellent contributors. In the late nineties, Basel II was on everyone's mind, especially the discussions around internal models. It was then that Operational Risk (OpRisk) was introduced as a new risk category. Several years later, we provided the first article [4] of the newly established Journal of Operational Risk precisely on the topic of (very) extreme events within OpRisk: little did we know back then that a decade later operational risk losses would pale many other banking losses on Wall Street. From today's point of view, the inclusion of OpRisk under Pillar I of Basel II may be seen as excellent foresightedness. Already then OpRisk was indeed perceived as an intrinsically import risk category. One further reason for its inclusion was that regulators wanted a compensation for the anticipated reduction in regulatory capital for MR and especially CR at a global industry-wide level. The above examples show the strong involvement we had with the early publication outlets of the Risk Waters Group, later to become part of Incisive Media. It is not accidental that around the same time, as a matter of fact on October 7, 1994, at the Department of Mathematics of ETH Zurich we founded, in collaboration with

the Swiss financial industry, RiskLab (see www.risklab.ch). Through J.P. Morgan and RiskMetrics, Value-at-Risk (VaR) was just introduced to the banking industry, and this in partial response to the famous 4:15 Weatherstone report. The then CEO of J.P. Morgan, Dennis Weatherstone, wanted (daily, by late afternoon, hence 4:15) an answer to the simple question: "How much could J.P. Morgan lose if tomorrow turns out to be a bad day?" In this context, at RiskLab, very early on we discussed the pros and cons of VaR-type risk measures. Topics like netting of derivative positions and internal model approaches were being pushed by the financial industry leading to interesting mathematical research problems. All this activity led to increased discussions between academics and practitioners, i.e. to an increased link between academia and the professionals working in financial institutions. RISK publications played a pivotal role in this interchange. At RiskLab, our aim was (and still is) to provide a discussion platform where relevant technical issues in the realm of QRM can be discussed openly; we coined the phrase "RiskLab offers a **precompetitive** discussion platform for QRM related problems." Some of these discussions led to scientific publications which are now part of standard QRM technology**.** Beyond the introduction of EVT within QRM, it is worth mentioning two more relevant examples: the work on models for dependence concepts between underlying random variables or risk factors going beyond linear correlation (i.e. copulas [5]) and the seminal work on coherence of risk measures [6]. Extended versions of these papers were published in research journals but practical summaries appeared early on in RISK Magazine, very much facilitating the dissemination of the fundamental research driven results in the industry. Both papers were also included in [3] making up a section on "Risk Measures and Extreme Value Theory." In these publications one also learns about the potential pitfalls - and hence limitations - associated with VaR as a risk measure and the advantages to assess risk by way of a coherent, in particular a sub-additive risk measure such as Expected Shortfall (ES). Hence the resulting push from within RiskLab for the adoption of ES as a regulatory risk measure instead of VaR. The debate about which risk measure (VaR or ES) to use in practice and for what specific purpose is still ongoing. The dialogue and interaction between professionals and academics has been further promoted and nurtured through the organization of a yearly Risk Day at the ETH RiskLab. Indeed, some of the aforementioned fundamental research and ideas were presented to the (local) financial industry at the kickoff event in 1998; see [7]. These early

publications, further development and results accumulated in subsequent years were summarized in the standard QRM textbooks [8] and [9]. The content of these books is supported and complemented through the webpage www.qrmtutorial.org where most R-program-based routines and algorithms used in the book can be found.

Next I will reminisce on the development of QRM ideas in finance and insurance, especially concerning internal models, from a current as well as future perspective.

Internal models: a first digression

Ever since internal models came onto the regulatory scene, the world of banking and insurance has witnessed "a survival of the fittest evolution" of analytic models across the various Pillar I product-type subcategories. Within the regulatory frameworks worldwide, e.g. the International Basel Accords (especially Basel II and Basel III) for banking and the European Solvency II and Swiss Solvency Test (SST) guidelines for insurance, larger (international) institutions were allowed, even encouraged to come up with internal models towards the calculation of regulatory capital, best fitting the institution's business profile and product range. The institutions themselves very much lobbied for the use of internal models claiming their superior risk-sensitivity. At the same time, they also aimed, quite naturally, for lower regulatory capital charges e.g. through diversification effects. The industry settled rather quickly on a fairly broad set of internal models to account for MR and CR, including models for standard products widely available and traded in the market. And whenever new products appeared, their introduction was followed, in a natural way, by further model developments and model selection processes. In the case of OpRisk, though many advanced models were introduced, this "evolution" or "race" to identifying a (most) appropriate or widely agreed upon internal model under Basel II's Pillar I, did not really work well. There are many reasons for this: data-scarcity, data-quality and overall problem-complexity, the extreme non-homogeneity between the various risk subclasses of OpRisk, e.g. ranging from internal fraud, over external events to legal risk, and finally the Basel II guideline that Pillar I regulatory capital for OpRisk had to be calculated using a yearly 99.9% VaR. Insurance regulation (e.g. SST) realized these fundamental issues early on,

whereas banking regulation only now comes to grips. Finally, the 2007-2009 financial crisis provoked a regulatory onslaught against the use of internal models that is still ongoing. The discussions around the appropriateness, effectiveness, robustness and reliability of internal models to manage risk contributed significantly to a backlash to quantitative (e.g. mathematical) models used within banking and insurance regulation. Adding to the lack of understanding and confusion (or perhaps trying to divert responsibility and sheer incompetence of real culprits intoxicated by a culture of excess, greed and irrational exuberance), a number of more popular publications achieved Hollywood style "level" rather than scientific seriousness like in the case of the Gaussian copula and how the underlying mathematical formula would have destroyed Wall Street; see [10]. At the time, even the more serious financial press jumped on the let's-blame-the-quants bandwagon; see [11]. The authors of these publications would have been well advised to first have read the thoughtful interview [12] given by Steven Shreve, about one year earlier, on the topic of (just) blaming quants for the financial crisis: "The quants know better than anyone how their models can fail. For banks, the only way to avoid a repetition of the current crisis is to measure and control all their risks, including the risk that their models give incorrect results. On the other hand, the surest way to repeat this disaster is to trust the models blindly while taking large-scale advantage of situations where they seem to provide trading strategies that would yield results too good to be true. Because this bridge will be rebuilt, **the way out of our present dilemma is not to blame the quants. We must instead hire good ones–and listen to them**", a statement still highly relevant today, especially concerning the importance of quantitative (i.e. internal) models. In fact, my own view in this respect was expressed already in 2001(!) in our early response [13] to the, at the time, new Basel II guidelines where we very clearly put our academic finger on the weaknesses of the new regulatory proposals and most strongly warned for possible disastrous consequences. Some of the weaknesses we highlighted included the neglect of endogeneity of risk, network vulnerability and systemic risk, possible procyclicality of the new guidelines, the data issue underlying OpRisk, the widespread (mis)use of rating agency AAA labels for complicated financial products, as well as some more technical suggestions on the use of risk measures (in particular supporting the use of Expected Shortfall instead of Value-at-Risk). On page 5 of [13] we summarized these concerns in the direct crude style that is

common to academics: “Reconsider before it is too late.” (Unfortunately, but certainly not surprising, about 5 years later it was too late!) This “Academic response” was officially submitted to the Basel Committee on Banking Supervision, a step I very much hope more of my academic colleagues would take. Whereas [13] received considerable attention and was widely read and cited to this day, its immediate influence on the Basel II proposals was minimal. With hindsight, we could and should have done more. The spot-on conclusions in [13] were based on findings gathered at a broadly attended conference on Basel II organized by the Financial Markets Group of the London School of Economics, and as such was not merely based on ivory-tower-thinking. Recognition for this and subsequent work came much later when Lord Adair Turner, then Chairman of the UK Financial Services Authority (FSA), invited me to attend a conference in London, 22 March 2010, organized by the FSA. More precisely the invitation request read: “I would be delighted if you could join us at this event to speak about the modelling of traded assets and its role in prudential regulation.” It is fair to say that some of the findings of [13] found their way, directly or indirectly, in future regulatory guidelines.

Some relevant questions

In the current discussion concerning quantitative (internal) models in finance and insurance, here are some questions which need answering:

(Q1) What precisely constitutes a model;

(Q2) What is the role of calibration and how can models be validated, for instance through statistical backtesting or, more precisely, through passing a series of appropriate controls based on criteria and tests that are part of a sound validation framework with respect to which models are evaluated and eventually accepted (or better formulated from a statistical hypothesis testing point of view, not rejected);

(Q3) Where and how do model uncertainty and model robustness enter;

(Q4) How to communicate how models are used within an institution, and how can findings from these models be reported to the outside world, e.g. in order to establish regulatory benchmarks and /or to comply with Pillar II guidelines;

(Q5) How to add stress testing components and more importantly, which stress tests to use;

(Q6) What are the practical limitations and potential shortcomings of “one model to rule them all”, and indeed of “one risk number (VaR or ES) to rule them all” and how do these possible concentrations have an impact on systemic risk;

(Q7) What are the conceptual and technical issues in going from a bank internal model landscape (or warehouse) to measures of solvency;

(Q8) Is there a need for balancing or combining standardized approaches with approaches based on internal models;

(Q9) What would be the consequences of the disallowance of the use of internal models for solvency purposes, and (for the moment) finally

(Q10) How should we view the internal versus standardized model dispute (and more broadly, the need for much more fundamental changes) in a business environment that is changing at an accelerated rate due to digitalization?

And some answers

Let me first comment on (Q9). It is important to realize that the discussion on the regulatory allowance or not of internal models mainly refers to solvency/capital calculations. Indeed - and this by definition itself - standardized models typically allow for a better (e.g. more congruent) comparison between financial institutions having a comparable risk or business profile. The word comparison is important here. More questionable is the lack of risk sensitivity to (especially adverse) market conditions. Concerning internal models used before and throughout the financial crisis, we have learned from regulatory studies across the wider banking industry that risk capital calculations for a given reference portfolio resulted in widely diverging capital numbers based on those models. As a consequence, risk capital numbers between banks became difficult to compare in an objective way. At the level of bank-internal risk management (including pricing and hedging at the product level) it is clear that internal models still remain the standard. In an ideal world, the models used for setting capital should not be too different from those used for internal risk management purposes. A too strong dislocation between “the two worlds of risk measurement” cannot be an ultimate goal.

Concerning (Q10), the changes alluded to here are mainly driven by Big-Data-Management (better referred to as Data Science) and Information Technology (IT). Buzzwords include Algorithmic Trading, High Frequency Finance, Neural Networks, Machine Learning, Blockchain Technology, Cryptocurrencies, Distributed Ledger, Smart Contracts and the numerous developments in the FinTech/FinReg universe. In particular, the Big-Data developments herald a move from “Know your client” to “Know your data”, a move I personally am not very comfortable with. Below, I will comment only on some of the underlying issues. Proper and adequate assessment of the advantages and disadvantages in employing a standard or an internal model depends on various considerations (scope and purpose). For this and other reasons, there is no universal framework with respect to which we can identify a clear winner between standard and internal models (see also my comment under (Q9)). Their appropriateness depends crucially on the context and thus so is the choice of which model to employ. Whatever solution one comes up with, a balance between risk sensitivity, simplicity and comparability must be the final goal. One thing is for sure, in order for an internal model to become a gold standard in this changing and highly competitive technological environment, it has to be **fully understood**; there is no room for black box magic. Models are tools that help us sharpening the questions that are being raised leading to a potential increase in understanding and to support decision taking. Thus far models do not take decisions, people do; perhaps we may be witnessing exactly now the transition from “thus far models do not” to “now models do”. For this reason, one has to be able to communicate the results from such internal models, as well as the underlying model assumptions, to a sufficiently wide audience in a clear, succinct and understandable way. Will companies and not just models seize the digital challenge and adapt? Clearly, the discussion about models and model development goes well beyond the realm of internal models, especially in view of a broad advance of AI-based technology in all aspects of daily life. Enrique Loubet (personal communication) summarized the human-machine-decision-process as follows: “It would be misleading and false to claim that models will ever be "taking decisions." Most certainly, models will indeed become progressively more sophisticated as various steps in model development and model interfaces that currently require man-assessed-choices and inputs (i.e. steps involving human interventions and decisions) are likely to be automated and hence integrated into

them. But this event **will** be a human decision on its own. That is to say, we cannot pretend to hide responsibility in an automated process. Humans, either as model developers, users of models or people taking decisions based on the output of models, are **interacting with all** aspects of the algorithms forming part of an ever increasing push for automation.“ As it stands, however, the current discussion concerning regulation for insurance and banking is dwarfed by the much wider, more political and societal debate of what constitutes the ideal (or at least better) financial architecture of the future. The final verses of Jorge Luis Borges’ poem Ajedrez/Chess come to mind:

… God moves the player, and he, the pieces.

Which God behind God starts the plot

of dust and time and dreams and agony?

In the context of the present discussion around internal models, we could say the following: “Internal models/quantitative models are embedded in corporate ones and these in societal, ethical, political and governmental ones. That is the complex and intertwined environment where we live, try to adapt and function in. We can try to narrow the discussion to a component of this whole only to realize that when going deeper we often can no longer remain in the isolated framework where we started.”

A new architecture for financial institutions and its regulation

It is not surprising that politicians, academics, regulators as well as the general public want to reign in the complexity of banking institutions; see e.g. the discussions around Limited Purpose Banking [14] and whether or not our bankers are indeed “wearing new clothes” [15]. At the same time, already by now “classical” insurance and finance has become highly technical and this trend is expected to sharpen at an ever accelerated rate. For insurance for instance, it surely would be a bad move to turn away from Market Consistent Valuation (MCV). Solvency I (i.e. so-called statutory) figures published by the Swiss regulator FINMA for life insurance companies headquartered in Switzerland hardly showed any effect of the financial crisis; so much for risk sensitivity! Insurance regulation is very much based on policyholder protection and solvency guidelines aiming for an “at arm’s length” possible transfer of business between two parties in times of

distress; at such moments in time MCV is absolutely crucial and properly worked out internal models (including results from well-chosen stress tests) offer key guidance here. This point of view is very much akin to the concept of a "living will" for banks. I deliberately used "classical" above; indeed, in a rapidly changing, increasingly technologically driven world, internal models together with the intellectual capacity within companies and regulatory bodies will prove to be critical. We (regulation, industry, and academia) need to be able to attract the best minds in order to face up to current and future challenges. Whereas today the regulatory horizon is typically a year (though intermediate checks do of course take place), new products as well as the ambient IT landscape will move from periodic (regulatory) oversight to a much more dynamic form of supervision. A statement to that effect was made by Nobuchika Mori, the Commissioner of the Financial Services Agency of Japan [16]: "The safety and soundness of a bank cannot be captured by a point-in-time assessment of its balance sheet alone. They are ensured through dynamic interactions between the bank and the markets, and affected by various elements in the entire economy." And further: "… we intend to move from a framework dominated by static regulation to that complemented by dynamic supervision." But also: "… the global regulatory community aspired to maintain financial stability and enable sustainable growth by providing banks with **incentives to enhance their risk management practices**, capital strategies and business models. The JFSA is hoping to explore the potential benefits of such an approach once again." The bold type for emphasis (which is mine) seems to hint at a possible rewarding of a proper use of internal models for risk management purposes. Indeed, early on in his speech, Nobuchika Mori made the comment that "The global regulatory community's preoccupation shifted from bettering risk management to enhancing capital adequacy. Less confidence is given to supervisory processes adapted to specific institutions and more hope is placed on the effectiveness of uniform rules. It is sometimes argued that the room for innovation in risk management can be abused by arbitrage and that regulators need to intervene deeper into banks' risk management processes." The point on regulatory arbitrage is well taken, but does not solely apply to the banking world. The reader is for instance encouraged to look at an example of shadow insurance which grew out of regulatory optimization/arbitrage; see [17], [18] and [19]. In view of such, and several other similar examples/constructions, it

does seem strange that a lot of current regulatory effort is aimed at moving away from internal-model-thinking.

Capital ratios and regulatory arbitrage

The calculation of regulatory capital constitutes the main battle field between regulators and financial institutions when it comes to either use internal or standardized models. Regulatory capital always constitutes a quotient with capital in the numerator and a measure of risk in the balance sheet in the denominator; reporting higher regulatory capital numbers is hence not only achieved through "reducing" the Risk Weighted Assets, say, in the denominator, but also by "increasing" capital in the numerator. Both words, "reducing" and "increasing", are in quotation marks as history has shown that a combination of clever financial engineering, creative accounting, and tax optimization may be misused to this effect. Banks as well as insurance companies deliver **products** like loans, client portfolios, risk hedges, alternative risk transfers, life and non-life insurance products, pension solutions … and hence a proper quality control of such products (like in the manufacturing industry) can be expected. Internal models become eminently important to assess and report the risk embedded in these products and to support the communication with clients of the potential benefits and risks associated to their transfers. A well-balanced point of view on the wider debate was given by Isabelle Vaillant, the Director of Regulation at the European Banking Authority (EBA). Her summary statement in [20] is that the Authority's "key goal" in Basel talks has been to defend a risk-sensitive capital framework, in particular: "Non-risk sensitive items should be a non-dominant part of the framework" and further "It is not a mystery that the EBA has been defending **models** with the idea that **risk sensitivity is crucial**". For the non-specialist, it is definitely useful to follow up on some concrete discussions between industry and regulation on the topic of internal models. For example, in the context of Market Risk, one place to start is EBA's December 14, 2015, document EBA/CP/2015/27 together with several industry responses, e.g. including ISDA's and afme's. EBA's final draft, submitted to the European Commission, came about on November 22, 2016 as EBA/RTS/2016/07, a 127-page document. A somewhat broader scope on the use of internal models is found in EBA's February 28, 2017 document, a "Guide for the Targeted Review of Internal Models (TRIM)" (155 pages). The following statement from the Foreword of the latter document is telling in connection to the ongoing

debate between industry and regulation: “The Targeted Review of Internal Models (TRIM) is aimed at enhancing the credibility and confirming the adequacy and appropriateness of approved Pillar I internal models permitted for use by Systemically Important Financial Institutions (SIFIs) when calculating own funds requirements. The category of SIFI institutions was introduced in the wake of the financial crisis in order to better safeguard the global financial system against the threat of systemic risk. As a major objective, TRIM focuses on the **reduction of unwarranted variability in risk-weighted assets (RWA)** driven by inappropriate modelling which takes advantage of the freedom granted by the current regulation.” I will come back to this statement when commenting on the so-called London Whale.

Models and the market

Concerning **models**, it is difficult to beat the often quoted statement “All models are wrong, some are useful” by the statistician George E.P. Box. The full statement from his 1976 publication [21], addressing the issue of parsimony, reads as follows: “Since all models are wrong the scientist cannot obtain a “correct” one by excessive elaboration. On the contrary following William of Occam he should seek an economical description of natural phenomena. Just as the ability to devise simple but evocative models is the signature of the great scientist so overelaboration and overparameterization is often the mark of mediocrity”, a statement to be appreciated in the current deluge of machine learning and neural network technology. Whereas in the above quote one can - and for the purpose of this discussion, one should - augment “natural” by “economic”, that broader interpretation has its consequences. First of all, most scientific models developed to describe natural phenomena are driven by incorporating sensible hypotheses from empirical observations. The scientific premises and predictions of these models being tested by conducting experiments which other scientists can repeat to corroborate or refute the results. Indeed, following Karl Popper, scientific models cannot be “validated,” they can only be “falsified” as soon as predictions of the model fail to match observations, in which case the models need to be reviewed and amended to extend their scope. The financial market is a dynamic environment with multitudes of feedbacks and thus, does not offer the possibility to “repeat” or run experiments under controlled conditions to assess models in a peer-reviewed way

as in science. For instance, in many ways we navigate in the dark when it comes to macroeconomics, a fact amply confirmed by the several financial crises and the reactions to these by regulators, politicians and central bankers. How would one repeat or even test in an experimental economics environment macroeconomic scenarios? One particular difficulty stems from the fact that there exists a strong feedback to the market from models and the financial products that are being introduced; in a way, they **define** "the market" (see [22] for a more philosophical discussion on this notion of "market"). Unlike in the case of most scientific models - except perhaps when dealing at small quantum scales - where the model and measuring devices do not distort the underlying physical phenomena in any significant way, financial models and active financial products directly impact and change the market, i.e. they define or affect the object they aim at describing. In finance, model development is driven by idealized assumptions on (a) the functioning of, and (b) the agents participating and interacting within a financial market, often mathematically glued together with (c) no-arbitrage arguments. It is not unrealistic to claim that, when it comes to the world of banking and finance, opacity reigns with little data to guide us. Non-linear effects appear in many, if not most of the financial derivatives, and intricate dependencies compound to make precise modeling a very hard task indeed.

A capital adequacy Ansatz, be it standard- or internal-model-based, always presupposes conditions on the underlying markets. These conditions are (or should be!) more obvious in the case of an internal model due to the explicit quantitative modeling. On the other hand, standardized approaches are also based on assumptions, at least involving the use of specific accounting rules. Browsing through the various standard models in Basel III, say, reveals a non-trivial amount of complexity and (often implicit) underlying assumptions. In both cases it is essential that assumptions are made explicit, clearly communicated to the various stakeholders involved and adhered to throughout, up to the actions based on the conclusions they lead to. As an example of such discussions, consider the important case of OpRisk. After several years of an Advanced Measurement Approach (AMA) Ansatz for the modeling of OpRisk, the Basel Committee has recently decided against the use of internal models while supporting a standardized approach instead. Over recent years it became clear that great variability towards the calculation of OpRisk regulatory capital resulted

from the different AMA models (and hence assumptions) used by larger international banks. The proposed standard model however also involves assumptions that need to be made explicit. For a detailed discussion on this, see [23] as well as the references in the latter paper to the relevant Basel Committee documents.

Regulatory arbitrage, rating agencies and product complexity

Whichever approach one takes, regulatory arbitrage is to be avoided, for which regulatory vigilance is needed. The unfortunate prime negative example was the development of the CDO and CDS markets leading up to the 2007-2009 financial crisis. All too often an AAA-label was most gladly (and passively) accepted without questioning the rating agencies' (internal) models used for arriving at this label (at this point, **reread** the statement on quants made by Shreve earlier in the paper; see also our 2001 (!) warning about these practices in [13]). Further, AAA ratings were typically interpreted as meaning risk-free, leading to massive buying and, more dangerously, to a bank-internal warehousing of such products. In doing so, banks, as well as some insurance companies, typically with the suffix FP (standing for Financial Products) to their name, violated their life-long raison d'être of providing maturity transformation and the acquisition and the selling of risk through **properly** diversified portfolios. As a consequence, by 2007, the financial industry as a whole was (akin to being) long an economic catastrophe bond waiting to be triggered by a macro event which, in the end, turned out to be a substantial decline in the American housing market ([24], [25]). The ensuing spillover of problems from Wall Street to Main Street is to this day present in our minds. Regulatory arbitrage between the banking book and the trading book allowed for a considerable reduction of internal-model-based risk capital charges. A further issue that surfaced throughout the crisis concerned the calculation of RWAs where minor model "corrections" led to considerable reductions in the risk capital (required to be) reported. An unsavory example of the latter is to be found in the case of the so-called London Whale; read for instance the story as told in Jacque's excellent book [26], especially the section on "The Art of Concealment", p. 300-304 (see also my comments above related to EBA's TRIM document). One may add to this list Lehman Brothers' legal-arbitrage of their pre-default leverage ratios using a REPO-105/108 accounting maneuver, as well as the mis-selling by several investment banks of opaque financial products. All too often, the

originators of such products were themselves vague, to say the least, about some of their products' societal value, nor did they always understand them fully themselves, and as a consequence were not able to communicate clearly to the eventual buyer the underlying risks. After the crisis, this fact led to numerous fines levied on the financial industry, fines that were then booked under the regulatory denominator of OpRisk and this within the category of legal risk; by late 2006, the cumulative legal fines to the banking industry amounted to more than 220 billion USD. In terms of market capitalization of individual (mainly investment) banks, values from 10-20%, in the case of Bank of America even 50%, were reported (see The Economist, 13 August, 2006). It is therefore no wonder that observers started questioning the ethical standards and lack of social responsibility of bankers. The current "attack" on internal models is just one consequential aspect. Another major issue, especially for smaller countries with several large, systemic international banks and insurance companies, is that the regulatory inspection of the internal models used at such institutions is very costly, i.e. putting an additional stress on capital and human resources. This is an observation difficult to bypass. It brings us back to the discussion on increased sophistication of models and over-complexity versus the actual relevance and value to society of many of the financial products that are produced and sold.

No doubt, quantitative models in financial risk management have brought considerable success to the economy as a whole and one cannot even start thinking of "getting rid of such tools" for trading, risk management and solvency purposes. And this even more so at a time where innovative alternative risk transfer solutions are increasingly in demand, e.g. in the world of environmental risk. Internal models ought to be integrated into the core processes of financial and insurance institutions. This is one of the guiding principles underlying the Pillar II approach within Basel II/III. In the realm of insurance regulation this corresponds to Own Risk and Solvency Assessment (ORSA, e.g. [27]). ORSA for instance stresses the fact that there always will be a fine balance between (internal) models and leadership, but concludes the obvious fact that "Models cannot replace leadership" ([27] p. 25). In regulatory practice, both for banking and insurance, forcing market participants to strongly limit the choice of models must be avoided. The Darwinian "best surviving model" often does not exist. And as such, model diversity is not necessarily bad, as long as these alternatives (and

their underlying assumptions and limits) are fully understood and communicated. An example that clearly shows some of the dangers in risk management lurking in the concentration on one specific model was abundantly made clear during the 1987 crash, in part due to the widespread use of VaR-centred risk management and of program trading based on portfolio insurance; see e.g. [28] for a 2007 view on that crash. Recent losses in the realm of algorithmic and computer-program-triggered trading raise similar concerns. In times of crises, “cash is king” and liquidity risk surrounding models often comes unpleasantly to the forefront causing everyone to panic and “run for the exit”; see [29].

Risk governance

Corporate governance ought to function in such a way that internal models are there to enhance the overall institution’s performance to the benefit of all stakeholders involved. It is nonsense to say (and here I quote an occasionally encountered criticism) that “internal models do not capture tail risk” or that they cannot handle “complex interdependencies”; of course they can, but such model features have to (and can) be included, often leading to higher capital charges, so be it! One unpleasant consequence from the regulatory drive away from the use of internal models for capital and solvency calculations is that regulators may face difficulty retaining their better quantitative people; such development cannot be healthy in the long run whatever the financial architecture is that will prevail in the future. Already now, the world of banking and insurance is not an easy one to fully grasp. Add to that existing complexity ingredients like interconnectedness (either social or technological), demographic changes, a major drive to just-in-time production and delivery (a stress in pursuit of higher efficiency and productivity at the expense of increasing vulnerability leading to business interruption as one of the main emerging risk categories facing the world of insurance), but also the whole drive towards a Big-Data-oriented society with almost continuous monitoring at all levels, and the ensuing vulnerability and threat of cyber-crime - to name just a few of the changes just around the corner - then it becomes abundantly clear that we need “all hands on deck.” Any good risk management department must contain a diversity of talent, skills, competences and experience. Throwing out quantitative skills and reducing capital adequacy considerations to the predominantly qualitative/standardized level is necessarily incomplete and cannot be the ultimate goal. As Albert Einstein is often quoted to

have said: “Everything should be made as simple as possible. But not simpler.” This is reminiscent to Occam’s razor or the Law of Parsimony we encountered in George Box’s earlier quote on models.

A sample of some research results

Independent from an immediate practical use, all the issues above (and many more) lead to fascinating research questions, questions which, ideally, are to be tackled through a close collaboration and constant dialogue between industry and academia, including guiding discussions with regulatory bodies worldwide. This brings us back to the spirit very much prevailing around the birth of RiskLab in 1994 and the early publications around RISK and the JoR. In these early publications it was made abundantly clear that VaR, standard deviation and linear correlation are misleading risk metrics (to say the least) when it comes to applications in markets where the stochastic behavior of the underlying risk drivers is “well beyond the bell-curve.” Already in [5] we formulated, somewhat tongue in cheek, **The First Fundamental Theorem of QRM** stating that within the world of elliptical models (e.g. the multivariate normal or multivariate Student-t) the three risk metrics above work fine; for a precise formulation, see for instance Theorem 8.28 in [9]. At the time, we however stressed the much more important **Second Fundamental “Theorem” of QRM**: in a non-elliptical world (i.e. in reality!) all the conclusions of the First Theorem fail, and depending on how far reality deviates from the realm of elliptical models, the extent of the failure may be significant. In a non-elliptical world, VaR is non-coherent, standard deviation becomes a questionable measure of risk and linear correlation is not able to accurately capture dependencies. Unlike the First Theorem, the Second “Theorem” as such is not one that is rigorously or precisely formulated. Nonetheless it summarizes numerous mathematical results from the world of Model Risk [30] and Robustness, which are fields of current research of considerable importance. In view of some of these results, it is deplorable that VaR is still predominantly used for risk management practice as the benchmark for managing financial risk; see [31]. VaR, as a high quantile of a Profit-and-Loss (P&L) distribution, is fine to report; actuaries and engineers have used this risk measure with great success for a very long time. In their language, this measure is referred to as a “once over a given time period return event.” Ample examples can be found in the reporting of earthquakes, floods and storm events. Statistical

estimation of these risk measures remains difficult, especially at very high quantile levels, long return periods. However, engineers would never start adding up such return-level-risk-measures, but that is exactly what the typical practitioner in the financial industry does on a daily (if not a minute) basis when VaR is applied to questions related to portfolio optimization, risk aggregation, diversification, and allocation. Then it just becomes the wrong choice of risk metric in any realistic market environment. If (and that is a big if) a risk metric is to be used as a P&L summary, then Expected Shortfall (ES) is far superior! Its general convexity properties (see [32]) allow e.g. for its use in risk aggregation and allocation applications. Equally important, by moving away from VaR-based to ES-based risk reporting, one moves from an "if" to a "what if" oriented risk management culture. The crucial point being that, for instance, as an answer to the question, raised in 2006, "What happens to our MBS portfolio if over the next two years, say, American house prices fall by 20%?", an "if" reaction would be "The probability of such an event is astronomically small, it will not happen!". On the other hand, in a "what if" discussion one simply asks "**What** are the consequences for our MBS portfolio **if** that happens?" If the answer to that "**what if**" question would be "We stand to lose several billion USD", then surely some managers or board members higher up in the hierarchy would (at least) raise their eyebrows. It is the enormous volume increase and warehousing of perceived riskless assets that created the financial crisis. The last example is based on actual facts; the real underlying example resulted in a 50 billion USD rescue plan. For more discussions, examples, but also warnings on these issues, see [9], or read some of the recent papers on my website www.math.ethz.ch/~embrechts.

Summary

Of course, I could have addressed many more aspects on the use and misuse of models in banking and insurance practice and regulation. Below I give a partial summary of the discussion so far, stressing a bit more one or two points:

(S1) I believe that Internal Models in the financial industry are here to stay, that they need to be well understood and carefully documented, and that they be properly challenged and critically calibrated. Institutions are well advised to methodologically catalogue them and check for consistency of their usage across different product and reporting lines, resulting in what one could refer to as an

internal-model-book. At the level of calibration, the difference between internal and standard models may be blurred. A standard model Ansatz may readily transform into a highly internal one at this stage.

(S2) Internal models play an important role for regulatory purposes, as long as their use is methodologically clear, scientifically sound and ethically correct.

(S3) I think it is important to understand changes in model-based capital values over certain time periods, instead of concentrating solely on their standalone values.

(S4) I welcome an added value resulting from combining standard risk measurement procedures with internal-model-based ones; surely big discrepancies need reporting and require explanation.

(S5) I strongly believe that modern IT technology may for the first time (after the proverbial "ATM technology"-statement by Paul Volcker in 2009) push financial institutions in the direction of a very different business architecture; see for instance [33] for some historical comments related to "50 years ATM." Indeed, the ATM was "born" on June 27, 1967 as Barclays Bank installed the first Cash Dispensing Machine (CDM) at its Enfield branch in North London. The author of [33], John Shepherd-Barron, is broadly accredited as being the inventor of the CDM. As it is often the case with new discoveries, related ideas for this technological development were prevalent at the same time. In [34] one can find a detailed analysis of why it took the financial industry such a long time to embrace the full power of modern IT, now referred to as FinTech. In [35], a blueprint is given on financial regulation in a FinTech environment. Finally, the more legal challenges are broadly discussed in [36] leading to RegTech, standing for "**Tech**nical solutions to **Reg**ulatory processes". As a consequence of all these IT driven developments, the notion of "internal model" may/will take on a whole different meaning in markets of the (not too distant) future. Just like the already in place robo-advisor is different from our bowler hat wearing banker Mister Banks in Mary Poppins.

(S6) The determination of regulatory capital has important consequences for the emergence and growth of shadow markets, as well as for the products (not) being offered; the latter may even be more the case when this capital is solely based on standard models. And finally,

(S7) Corporate governance ought to function in such a way that internal models are there to enhance the overall industry's performance to the better of all stakeholders involved.

As an appendix to the summary statements (S1) – (S7), I like to quote from the latest Risk.net publication [37] which arrived on my desk just before I mailed my article to JoR, "ready for printing". Its opening sentence sets the scene: "Banks are doubling down on the use of machine learning techniques for model validation in the face of regulatory skepticism over 'black box' methods". One bank representative is quoted as follows: "Machine learning has proven particularly useful for **validating the models built for regulatory stress tests**, such as the Federal Reserve's Comprehensive Capital Analysis and Review." On the other hand, from the regulatory side one learns that: "Still, regulators remain wary about the use of machine learning in bank models. The Fed **warned against using machine learning to assess contagion risk** in model networks, saying these methods lack transparency and might obscure the true nature of banks' vulnerabilities." – Be that as it may, the battle lines are drawn, and Darwinian evolution in the internal model universe will run its course!

As I already made clear at the beginning of the paper, I used Charles Darwin's notions on natural selection and evolution in a somewhat light-hearted way. To conclude however, in a more serious vein on the "survival chances" of good internal models, it seems just appropriate to at least reflect more carefully on one of Darwin's famous quotes: "It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is most adaptable to change". One thing is for sure: there will be a lot of changes (in modeling and beyond) affecting and perhaps reshaping the world of banking and insurance. Industry, regulation and all the actors actively involved better shape up for the change "just around the corner".

Epilogue: 9/11

I started this article with some personal historical notes. I like to finish on a very personal one related to September 11, 2001. Earlier that year, I received an invitation to give a talk at the inaugural Waters Financial Technology Congress to be held in the morning of September 11, 2001, in the Windows on the World premises on the 106th and 107th floors of the North Tower (Building One) of the

World Trade Center in Manhattan, New York. Due to a clash with other commitments at the time, I reluctantly declined. I was already booked to give a talk on September 7 at the 11th International AFIR Colloquium in Toronto on the topic of "Bounds on VaR for general functions of dependent risks: the actuarial approach" and needed to be back in Switzerland the week of that fatal day. At 8:46 AM that morning of September 11, Flight 11 collided with the North Tower between the floors 93 and 99. All 65 conference participants already present and 16 staff members of the Risk Waters Group died; see [38] for a very personal account of these events by Peter Field, the founder of Risk Magazine. I like to dedicate this paper to the memory of these victims.

ACNOWLEDGMENTS

This paper grew out of discussions I had on the topic of internal models, and this over several years, with academics, practitioners and regulators. Early versions of my thoughts were presented at some international conferences. I take pleasure in thanking participants to these events. In particular, I like to thank Enrique Loubet for many thoughtful comments on an earlier version of the paper; his careful and critical reading very much convinced me of the importance of "critical thinkers" in our field. Needless to say that I solely am responsible for the accuracy of the statements made throughout the paper.

References:

[1] P. Embrechts, S. Resnick, G. Samorodnitsky (1998) Living on the edge. RISK, January 1998, **11(1)**, 96-100.

[2] P. Embrechts, C. Klueppelberg, T. Mikosch (1997) Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin.

[3] P. Embrechts, Editor (2000) Extremes and Integrated Risk Management. RISK Books, Risk Waters Group, London.

[4] J. Neslehova, P. Embrechts, V. Chavez-Demoulin (2006) Infinite mean models and the LDA for operational risk. J. Operational Risk, **1(1)**, 3-25.

[5] P. Embrechts, A.J. McNeil, D. Straumann (1999) Correlation: pitfalls and alternatives. RISK, May 1999, **12(5)**, 69-71.

[6] Ph. Artzner, F. Delbaen, J.-M. Eber, D. Heath (1997) Thinking coherently. RISK, November 1997, **11**, 68-71.

[7] Risk Day at ETH RiskLab. See www.math.ethz.ch/imsf/events/events.html www.math.ethz.ch/finance/ETH_Risk_Day.html (for 1998).

[8] A.J. McNeil, R. Frey, P. Embrechts (2005) Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Princeton, NJ.

[9] A.J. McNeil, R. Frey, P. Embrechts (2015) Quantitative Risk Management: Concepts, Techniques and Tools. Second Edition. Princeton University Press, Princeton, NJ.

[10] F. Salmon (2009) Recipe for disaster: the formula that killed Wall Street. Wired Magazine, February 23, 2009.

[11] S. Jones (2009) The formula that felled Wall St. Financial Times, April 24, 2009.

[12] S. Shreve (2008) Don't blame the quants. Forbes, October 8, 2008.

[13] J. Daníelsson, P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, H.S. Shin (2001) An academic response to Basel II. Special Paper 130, Financial Markets Group, London School of Economics.

[14] C. Chamley, L.J. Kotlikoff, H. Polemarchakis (2012) Limited purpose banking – Moving from "trust me" to "show me" banking. January 2012 preprint, Boston University.

[15] A. Admati, M. Hellwig (2014) The Bankers' New Clothes: What's Wrong with Banking and What to Do about It. Updated Edition. Princeton University Press, Princeton, NJ.

[16] N. Mori (2016) From static regulation to dynamic supervision. Keynote address, 31st Annual General Meeting of the International Swap and Derivatives Association, Tokyo, April 13, 2016.

[17] R.S.J. Koijen, M. Jogo (2016) Shadow insurance. Econometrica, 84(3), 1265-1287.

[18] R.J. Foley (2016) Iowa at center of debate over 'shadow insurance' deals. Associated Press, The San Diego Union-Tribune, August 29, 2016.

[19] B.F. Hepfer, J.H. Wilde, R.J. Wilson (2017) Taking shadow insurance out of the shadows: regulatory arbitrage, taxes, and capital. Mays Business School Research Paper No. 2836215.

[20] D. Wood (2017) Q&A: EBA's Vaillant on Basel IV, FRTB and CVA. Risk.net, April 4, 2017.

[21] G.E.P. Box (1976) Science and Statistics. J. American Statistical Association, **71(356)**, 791-799.

[22] E. Ayache (2010) The Blank Swan: The End of Probability. Wiley, Chichester.

[23] G.W. Peters, P.V. Shevchenko, B. Hassani, A. Chapelle (2016). Should the advanced measurement approach be replaced with the standardized measurement approach for Operational Risk? Journal of Operational Risk, **11(3)**, 1-49.

[24] V.V. Acharya, T. Cooley, M. Richardson, I. Walter (2009) Manufacturing tail risk: a perspective on the financial crisis of 2007-2009. Foundations and Trends in Finance, **4(4)**, 247-325.

[25] J.D. Coval, J.W. Jurek, E. Stafford (2009) Economic catastrophe bonds. American Economic Review, **99(3)**, 628-666.

[26] L.L. Jacque (2010) Global Derivative Debacles: From Theory to Malpractice. Second Edition. World Scientific, Singapore.

[27] G. Bernardino, EIOPA (2011) ORSA – The heart of Solvency II. Presentation at the Groupe Consultatif Summer School, Lisbon, 25 May 2011.

[28] J. Authers (2007) The anatomy of a crash: What the market upheavals of 1987 say about today. Financial Times, October 18, 2007.

[29] L.H. Pedersen (2009) When everyone runs for the exit. International J. Central Banking, **5(4)**, 177-199.

[30] M. Morini (2011) Understanding and Managing Model Risk: A Practical Guide for Quants, Traders and Validators. Wiley, New York.

[31] P. Jorion (2006) Value at Risk: The New Benchmark for Managing Financial Risk, Third Edition. Wiley, New York.

[32] P. Embrechts, R. Wang (2015) Seven proofs for the subadditivity of Expected Shortfall. Dependence Modeling, **3(1)**, 126 -140.

[33] J. Shepherd-Barron (2017) Meet the true star of financial innovation – the humble ATM. Financial Times, June 22, 2017.

[34] T. Philippon (2017) The FinTech opportunity. BIS Working Papers No. 655, August 2017. Bank for International Settlements, Basel.

[35] A.W. Lo (2016) Moore's Law vs. Murphy's Law in the financial system: who's winning? BIS Working Papers No. 564, May 2016. Bank for International Settlements, Basel.

[36] D.W. Arner, J. Barberis, R.P. Buckley (2017) FinTech, RegTech, and the reconceptualization of financial regulation. Northwestern J. International Law & Business, **37(3)**, 371 – 413.

[37] T. Osborn (2017) Banks tout machine learning amid regulatory concerns. Risk.net, September, 2017.

[38] P. Field (2002) Remembering September 11, the day I'll never forget. Risk.net, September 2, 2002.

Key messages

(KM1) It is important to have a historical perspective on some of the key developments in Quantitative Risk Management.

(KM2) The current challenges facing the financial industry ask for strong quantitative skills.

(KM3) Best practice risk management will always lead to a balanced combination of quantitative and qualitative skills, and of internal and standard models.

(KM4) The best (i.e. Darwinian) surviving models will be those that are able to adapt to a changing societal (including economic and technological) environment.

(KM5) The discussion around internal versus standardized models will take on a whole different meaning in a FinTech/RegTech driven world of finance and insurance.

Short CV

Since 1989, Paul Embrechts is Professor of (Insurance) Mathematics at the ETH Zurich, director of RiskLab and, since 2009, also Senior Chair of the Swiss Finance Institute. Besides numerous academic distinctions, he holds an Honorary Doctorate from the Universities of Waterloo, Heriot-Watt, Louvain, and City, the University of London. Previous academic positions were KU Leuven, the University of Limburg, Imperial College London and the London School of Economics. Dr Embrechts served as an independent director on the boards of companies in banking and insurance and co-authored the influential books "Modelling Extremal Events for Insurance and Finance", Springer, 1997, and "Quantitative Risk Management: Concepts, Techniques and Tools", Princeton University Press, 2005/2015. His extensive research has been published in leading international scientific journals; he is also a much demanded speaker at international conferences and events in academia, industry and for regulatory authorities. For a full CV, see http://www.math.ethz.ch/~embrechts