Global & Local Economic Review

Volume 22 No. 2 2018

Edited by
Nicola Mattoscio

Bruno de Finetti and the 'probability' of pursuing new economic knowledge

Vol. 22 No. 2
Global & Local Economic Review 2018
McGraw-Hill Education

€ 13,00 (i.i.)
www.mheducation.it
ISSN: 1722-4241

POLICY FOR THE SUBMISSION OF MANUSCRIPTS TO BE PUBLISHED IN GLOBAL & LOCAL ECONOMIC REVIEW

1. PRELIMINARY INSTRUCTIONS

For a manuscript to be accepted for publication in the Review, authors are required to abide by the following publication rules. In particular, authors must state that:

- the manuscript, or relevant parts of the manuscripts, was not previously published;
- the manuscript is not simultaneously under review by any other journal;
- the manuscript will be not submitted elsewhere before the final outcome of peer-reviewing process by the Scientific Board.

2. MANUSCRIPTS SUBMISSION

Manuscripts must be submitted to the Editorial Secretary by e-mail: gler@fondazionepescarabruzzo.it.

Manuscripts must be submitted following the instructions for authors (see http://www.gler.it/) and should include the main text, any footnotes, references, an Abstract not longer than 300 words, and one or more JEL codes available at https://www.aeaweb.org/econlit/jelCodes.php.

It is the author's responsibility to ensure that manuscripts are written in clear and comprehensible English.

Author's institutional affiliation, telephone number, and the e-mail address should be placed on the first page of the manuscript (the same applies to the corresponding author in case of co-authorship).

3. REVIEWING PROCESS

The manuscript is reviewed by the Scientific Board and, if judged suitable for publication in the Review, is then sent to three referees for a rigorous blind peer-reviewing process.

- The initial decision by the Scientific Board takes 2 weeks.
- Authors will be informed of the review decision and all relevant comments by the three referees within approximately a month.

Based on the referees' recommendations, the Scientific Board then decides whether the manuscript should be accepted as it is, revised and resubmitted, or rejected. In any case, authors will be provided with constructive feedback.

In case of acceptance, the publication process will begin by assigning the manuscript to a number and a year of the Review.

Proof corrections and the printing process takes approximately one month.

4. MISCELLANEOUS

In the event it becomes necessary in the future, authors submitting manuscripts to the Review are expected to be available for contributing as referees. The articles published in Global & Local Economic Review are listed in EconLit and eJEL bibliographies (American Economic Association).

// Global & Local Economic Review

Honorary Editor - Kumaraswamy V. Velupillai
Editor-in-chief - Nicola Mattoscio, University of Chieti-Pescara
Associate Editor - Tonio Di Battista, The Italian Statistics Society, Rome

«Global & Local Economic Review» is included in JEL on CD, e-JEL and Econlit,
the electronic indexing and abstracting service
of the American Economic Association

Editorial Offices: Corso Umberto I, n. 83 - 65122 Pescara
Tel. +39 0854219109 – Fax +39 0854219380
Website: www.gler.it
E-mail: gler@fondazionepescarabruzzo.it

Senior Portfolio Manager B&E: Marta Colnago
Portfolio Manager Custom: Daniele Bonanno
Produzione: Victoria Garcia Daganzo
Stampa: Logo srl, Borgoricco (PD)

McGraw-Hill Education (Italy), 2018
ISSN (print) 1722-4241 - ISSN (online) 1974-5125

A SPECIAL ISSUE OF

GLOBAL AND LOCAL ECONOMIC REVIEW

A Tribute to Bruno de Finetti
(Italian Scientist, Innsbruck 1906 – Rome 1985)

Bruno de Finetti and the 'probability' of pursuing new economic knowledge

Acknowledgements

The Editors and the members of the Editorial Board of G&LER are deeply grateful to Edgardo Bucciarelli and Donatella Furia for their valuable research work dedicated to this special issue.

Global & Local Economic Review

Volume 22 No. 2 **2018**

TABLE OF CONTENTS

Nicola Mattoscio
Bruno de Finetti's modernity: An introduction Pag. 1

Giuseppe Amari, Fulvia de Finetti
The logic of the 'uncertain' and Bruno de Finetti's civil mathematics " 9

Flavio Pressacco, Paolo Serafini, Laura Ziani
De Finetti and Markowitz mean variance approach to reinsurance and portfolio selection problems: A comparison " 27

Frank Lad, Giuseppe Sanfilippo
Scoring alternative forecast distributions: Completing the Kullback distance complex " 63

Francesco Carlucci
International monetary shocks and currency movements " 91

Giordano Bruno
The systemic thinking of Bruno de Finetti ″ 113

Simone Enrico Casellina, Giuseppe Pandolfo
Probability (of default) does not exist ″ 133

G. & L. E. R., Vol. 22 No. 2, 2018

Nicola Mattoscio*

BRUNO DE FINETTI'S MODERNITY: AN INTRODUCTION

This special issue is dedicated to the memory of Bruno de Finetti in recognition of his seminal research and teaching, his mentoring of generations of young researchers, and his undisputed leadership in the field of quantitative methods and even beyond. The figure of Bruno de Finetti, indeed, emerges in the context of Italian mathematics from the early decades of the Twentieth century both for the breadth of knowledge and the brilliant originality of his pioneering contributions, and for his disciplinary and interdisciplinary insights – as in the case of the probability calculus and the financial and actuarial mathematics – which have expanded the boundaries of mathematical knowledge of his and our time, to say the least.

Bruno de Finetti, undeniably, was certainly considered and is still known worldwide as a mathematician, or rather a leading mathematician, especially for being one of the founding fathers of subjective probability theory and a distinguished scholar of actuarial science. In particular, his fundamental contributions to the foundation of the theory of probability based on subjectivism made subjectivism itself a respectable sort of doctrine for a scientist to maintain. It is worth noting that these contributions emerged only a few years after the publication of Frank P. Ramsey's *Truth and Probability* of 1931 [1926], this despite he had neither known nor read Ramsey until 1937, that is, since de Finetti (1931), then reaffirmed in de Finetti (1937).

Much less known, even among some members of the scientific community, are his works – equally seminal – in the fields of economics, finance, and decision-making. Although de Finetti's fundamental contribution to economics lies in his concepts of probability, these lesser-known works allowed him to achieve some extremely important results long before than those scholars to whom these results have been traditionally attributed. To mention one example, de Finetti formulated the

* Editor-in-chief of Global & Local Economic Review, University of Chieti-Pescara, Italy, Tel.: +39-085-4537583. *E-mail address*: nicola.mattoscio@unich.it

notion of absolute risk aversion in 1952, anticipating Kenneth J. Arrow and John W. Pratt of more than one decade. He used this notion especially in connection with the risk premium in the case of betting and, furthermore, discussed the particular case of constant absolute risk aversion (see de Finetti, 1952). He also wrote anticipatory works on adverse selection (1938), martingales (1939), the optimal dividend distribution method (1957), and the Samuelson's model on interest rates with the choice between consumption and investment (1956).

The purpose of this special issue, however, is not merely reminiscent of Bruno de Finetti's vast body of work. The primary purpose is to reaffirm an analytic approach and its actual and potential application to different research fields, as well as to practical matters, which might be very useful to clarify some issues made even more cogent by the ongoing crises of the economy and economic theory. The methodological evolution of the economic discipline over the past decades, and more specifically the methods and content that are still characterising the debate on the subject today, suggest that it is particularly useful to recall Bruno de Finetti's general scientific approach and his specific contributions to economics and the economy as a whole (see also Gillies & Ietto-Gillies, 1987, 1991; Lunghini, 2006; Pressacco, 2006a; and Pizzuti, 2016).

These contributions, not without solid scientific backing, have generated further research questions and have helped to clarify a number of relevant economic issues with naturalness and analytical accuracy, which have been – and unfortunately are still very often – treated with rhetorical, unnecessarily sophisticated, and misleading methods in the debate mentioned above. It is also notable that these contributions are expressed especially within the welfare economics tradition.

In the 1967 Handbook of Insurance Economics, indeed, he wrote: "*The ways in which we could try to achieve greater social justice are complex and arduous. […] It seems that it should be the fundamental problem, the main problem, the most current problem, for economics scholars. As far as can be seen, why don't they care?*" (de Finetti, 2015, [1967, ch. VI], p. 177, Editor's translation). With this in mind, de Finetti himself put forward two possible rationales: "*First rationale: the nature of these problems is not congenial to the prevailing mindset among economists and to the mathematical-conceptual toolset they prefer. The most appropriate tools are in fact (in my opinion) those of 'welfare economics' […] which do not seem to be too popular (at least today)*" (ibid., ch. VI, pp. 177–178, Editor's translation).

For that matter, the second rationale lies in the fact that a practical study (*i.e.*, organizational and administrative sciences) would have been necessary to complement the abstract and theoretical approach of welfare economics and, therein, de Finetti's view of economics appears somewhat similar to that of other eminent scholars of his time such as Herbert A. Simon and Federico Caffè, although they may seem 'distant' from him – and 'different' from each other – but with both of whom de Finetti shared a sound conception of *utopia* (de Finetti, 1973).

Despite ups and downs, and the analytical rigor, de Finetti's economic writings are most frequently characterised by a common denominator represented by a great moral and methodological concern towards political and social change. He believes, indeed, that it is necessary "*to completely free oneself from the suggestion of what is, otherwise economics cannot become a science, but, at most, it can be developed as a series of good techniques to juggle many practical tasks.*" (de Finetti, 1935, p. 234, Editor's translation). According to de Finetti, therefore, economic research must play an active and constructive role, studying new organisational forms and new methodologies with open-mindedness and freedom from any conditioning in order to better assess where both limits and opportunities lie in the distribution of resources for efficiency and equity. In keeping with Hartog (1978), it is worth mentioning a general point here: during the so-called period of 'political economy', the relationship between economics and welfare remained below the surface, not because it was satisfactory resolved, but because the delimitation between economics and welfare was to a large extent blurred, while a sense of individuality prevailed among many economists.

At the time, there was not yet the maturity needed to posit the problem clearly and definitively. All this, apparently, was not (and still is not?) in the repertoire and the heartstrings of the vast majority of economists. The naïve and romantic phase in the relationship between economics and welfare lasted deep into the Twentieth century. Continuing with Hartog (1978), it can be argued that this phase came definitively to an end in 1932, "*when the sharp knife of Robbins carved a deep edge between politics as [the] chooser of ends and economics as analysis of means on the basis of given ends*." (Hartog, 1978, p. 456, square brackets added).

Two decades later, during the fifties, de Finetti implicitly reiterated this concept by focusing on the importance of methodological appropriateness of models used by economists and, specifically, warned of the widespread misuse of mathematics in the analysis of social insurance (see McCall, 1991; Pressacco, 2006b; Velupillai, 2015; and Pizzuti, 2016). Not surprisingly, in his 1956 report to the Conference on Actuarial and Statistical Issues on Social Security he addressed a few words of admonition (and, implicitly, of exhortation) concerning the main implications for (welfare) economics as a discipline: "*With regard to social security matters involving the entire community, the inadequacy of approaches based on purely formalistic aspects of mathematics was confirmed in all respects, without understanding in depth and without rethinking the demographic-and-economic bedrock, as well as without evaluating them in relation to the possible objectives and ideas of an ethical, political, and social nature. […] My impression is that [economists] are generally too entangled in the development of technical and financial superstructures to give the organisation of social security a more or less insurance role, and losing sight of the [main focus] and the question at issue.*" (de Finetti, 1969, pp. 281, and 284, Editor's translation, square brackets added).

This admonition, among others, preserves de Finetti's search for subtle connections among diverse phenomena, and is to be considered particularly significant if we reflect above all on the fact that it was advanced by a mathematician to the community of (welfare) economists already in the fifties of the last century. Speaking of which, whilst standard economic theory suggests a kind of universal, harmonious, and absolute logic to be applied both to economic phenomena and to human behaviour, its exacerbated focus on techniques – often as an end in itself and to the detriment of the foundations of economics – has raised growing concern particularly in the last four decades.

Thus, a fairly general conclusion, although not universal, that seems to emerge from this discussion is that the incorporation of a methodological and, therefore, paradigmatic approach to economic theory by standard economics has had and is still having negative consequences not only in terms of explanatory capacity of economic reality, but also and arguably more importantly in terms of people's welfare and well-being, up to the recent economic and financial crisis that affected global economy and has created a huge pressure on the responsibilities of economists and economics.

In all fairness, all things considered, if on one hand the main contribution offered to economic theory by Bruno de Finetti is the same to which he owes his reputation as a great mathematician (*i.e.*, his concept of subjective probability – not sufficiently assimilated and welded into one whole by many economists), on the other hand his contributions bearing a more or less direct relation to issues of an economic nature demonstrate a continued focus on the analytical aspects of economic theory, as well as on the epistemological, methodological, and political problems that come with it.

In the pages that follow, de Finetti's methodological approach together with resorting to much of his scientific legacy will be recalled. Some authors mentioned below were his students, and all the contributions refer to, or start from, de Finetti's very foundation and core ethos, drawing on his writings and his ideas. More specifically, the lead article in the issue is that by Bruno de Finetti's daughter, Fulvia, who wrote an interesting article with Giuseppe Amari concerning the logic of the 'uncertain' and de Finetti's civil mathematics. The peculiarity of this first contribution is the main text purposely published in Italian in the wake of deFinettian tradition.

The second article, written by Flavio Pressacco, Paolo Serafini, and Laura Ziani, emphasises the importance of comparing de Finetti and Markowitz mean-variance approach to reinsurance and portfolio selection problems. In the third article, Frank Lad and Giuseppe Sanfilippo develop two surprising new results regarding the use of proper scoring rules for evaluating the predictive quality of two alternative sequential forecast distributions, thus completing the Kullback distance complex.

The author of the fourth article, Francesco Carlucci, builds a vector autoregressive model linking the economies of the main monetary areas in order to consider international monetary shocks and currency movements, and to analyse the effects produced by monetary policies. Giordano Bruno is the author of the fifth article, in which he points out, methodologically, the systemic thinking of de Finetti. Simone Casellina and Giuseppe Pandolfo conclude the issue by delivering a *sui generis* written speech on simplicity and comparability of the risk measures and, in so doing, the two authors put the conceptual basis for a consistent discussion about risk measurement in banking and insurance environments.

In conclusion, this compact – yet intellectually challenging – collection in honour of Bruno de Finetti is a tribute to the originality, scientific rigour, and relevance to the development of science shown by a scientist of high moral character and of impeccable intellectual integrity, who have permitted succeeding generations of scholars to stand on his shoulders. This introductory editorial has significantly benefitted from the constructive criticisms of Edgardo Bucciarelli: I have endeavoured to write and revise this introduction to meet his most pertinent suggestions and insightful comments. Last but not least, this special issue of Global & Local Economic Review on Bruno de Finetti would not have been possible without the enthusiastic help and support of its Editorial Board.

References

de Finetti, B. (1931). Sul Significato Soggettivo della Probabilità. *Fundamenta Mathematicae*, 17, 298–329.

de Finetti, B. (1935). Vilfredo Pareto di Fronte ai suoi Critici Odierni. Nuovi Studi di Diritto, *Economia e Politica*, VIII(4-6), 225–244.

de Finetti, B. [1937] (1964). Foresight: Its Logical Laws, its Subjective Sources. *Annales de l'Institut Henri Poincaré*, Kybrug H.E. Jr. (trans.), reprinted in Henry E. Kyburg H.E. Jr. and Smokler H.E. (eds.), Studies in Subjective Probability, 7, 93–158, Kybrug H.E. Jr. (trans.).

de Finetti, B. (1938). Ai Margini del Dominio della Matematica nei Problemi di Assicurazione. *Assicurazioni*, 5(1), 365–382.

de Finetti, B. (1939). La Teoria del Rischio e il Problema della "Rovina dei Giocatori". *Giornale Istituto Italiano Attuari*, 10, 41–51.

de Finetti, B. (1957). *Lezioni di Matematica Attuariale*. Roma: Edizioni Ricerche.

de Finetti, B. (1969). *Un Matematico e l'Economia.* Introduzione di Siro Lombardini. Milano: Franco Angeli.

de Finetti, B. (1973). L'utopia come Presupposto Necessario per Ogni Impostazione Significativa della Scienza Economica. in B. de Finetti (ed.), *Requisiti per un Sistema Economico Accettabile in Relazione alle Esigenze della Collettività*, Milano: Franco Angeli, pp. 13–76.

de Finetti, B. (2015) [1967]. *Un Matematico tra Utopia e Riformismo*. In G. Amari & F. de Finetti (eds.), Roma: Ediesse.

Gillies, D.A., & Ietto-Gillies, G. (1987). Probability and Economics in the Works of Bruno de Finetti. *Economia Internazionale*, 40(3–4), 192–209.

Gillies, D.A., & Ietto-Gillies, G. (1991). Intersubjective Probability and Economics. *Review of Political Economy*, 3(4), 393–417.

Hartog, F. (1978). Hennipman on Economic Policy and Welfare Economics. *De Economist*, 126(4), 456–473.

Lunghini, G. (2006). *Bruno de Finetti and Economic Theory*. International Symposium: Bruno de Finetti Centenary Conference, Department of Mathematics, University of Rome, La Sapienza, 15–17 Novembre 2006.

McCall, J. J. (1991). *Vico and de Finetti: An Economic Perspective*. Working Paper no. 630, Department of Economics, University of California, Los Angeles (UCLA).

Pizzuti, F. R. (2016). L'Attualità di Bruno de Finetti nella Crisi in Corso dell'Economia e della Teoria Economica. *Moneta e Credito*, 69(275), 253–267.

Pressacco, F. (2006a). *Bruno de Finetti Hero of the two Worlds: (Applied) Mathematician and (Quantitative) Economist*. International Symposium: Bruno de Finetti Centenary Conference, Department of Mathematics, University of Rome, La Sapienza, 15–17 Novembre 2006.

Pressacco, F. (2006b). The Interaction between Economics and Mathematics in de Finetti's Thought and its Relevance in Finance, Decision Theory and Actuarial Sciences. *Giornale Istituto Italiano Attuari*, 69(1/2), 7–32.

Ramsey, F. P. (1931) [1926]. Truth and Probability, in D. H. Mellor (ed.), *Foundations: Essays in Philosophy, Logic, Mathematics and Economics* (1931; 1978), chapter 3, pp. 58−100. London: Routledge & Kegan Paul.

Robbins, L. (1932). *An essay on the nature and significance of economic science*. London: McMillan.

Velupillai, K. V. (2015). de Finetti's Theory of Probability and its Jaynesian Critique. *Economia Politica*, 32(1), 85–95.

G. & L. E. R., Vol. 22 No. 2, 2018

Giuseppe Amari * - Fulvia de Finetti **

THE LOGIC OF THE 'UNCERTAIN' AND BRUNO DE FINETTI'S CIVIL MATHEMATICS

Abstract

Bruno de Finetti's conception of mathematics and even more of subjective probability, of which he was one of the founding fathers, represent a research program of great and fruitful unitary relevance in the fields of logic, ethics and aesthetics. They are centered on the "logic of the uncertain" by which we proceed from "probable to probable", on the ethics of sincerity in the courage of the opinions, in the aesthetics of creative fantasy in response to the constraints of nature. The result is a radical, civil and reformist commitment always pursued by de Finetti which does not exclude the courage of utopia – not to be set aside a priori – but whose realization must always be carefully verified.

JEL CLASSIFICATION: A12, B23, B31, C02.
KEYWORDS: Coherence, Universal Welfare, Ecology.

* Fondazione Giuseppe Di Vittorio, via Gaetano Donizetti, n. 7/b, Rome, Italy. Phone: +39-06-85356715. *E-mail address*: g.amari@fdv.cgil.it.
** Fulvia de Finetti, piazza Filattiera, n. 49, Rome, Italy. *E-mail address*: fdfdefinetti@gmail.com.

LA LOGICA DELL'INCERTO E LA MATEMATICA CIVILE DI BRUNO DE FINETTI

Abstract

La concezione della matematica di Bruno de Finetti e ancor di più della probabilità soggettiva, di cui fu uno dei padri fondatori, rappresentano un programma di ricerca di grande e feconda rilevanza unitaria nei campi della logica, dell'etica e dell'estetica. Della "logica dell'incerto" in base alla quale procediamo "dal probabile al probabile", dell'etica della sincerità nel coraggio delle opinioni, dell'estetica della fantasia creatrice in risposta ai vincoli della natura. Ne deriva un radicale impegno civile e riformatore sempre perseguito da de Finetti che non esclude il coraggio dell'utopia – da non accantonare a priori – ma la cui realizzazione va sempre verificata con attenzione.

JEL CLASSIFICATION: A12, B23, B31, C02.
KEYWORDS: coerenza, welfare universale, ecologia.

Occorre pensare in termini di utopia,
perché ritenere di poter affrontare efficacemente i problemi
in maniera diversa
è ridicola utopia.
(Bruno de Finetti, 2015)

1. Coerenza tra la sua visione della scienza e il suo impegno civile

Che il nostro grande matematico nella sua vita abbia sempre perseguito un dedicato impegno civile, ne è esemplare testimonianza la sua intensissima biografia che qui non è certo il caso di ricordare[1] (de Finetti, Nicotra 2008).

Basti accennare almeno ad un episodio; quando fu arrestato per aver difeso, come direttore responsabile di un giornale del Partito Radicale del suo amico Marco Pannella, l'obiezione di coscienza per il servizio militare.

E dare la parola allo stesso Bruno de Finetti (di seguito BdF) in merito ai suoi dichiarati presupposti di valore:

"Per venire al ' chi sono ? ', la prima cosa che mi sembra di dover dire, come punto di partenza, è che di me stesso, come persona qualunque, m'importa assai meno che di ciò che attiene al benessere collettivo, all'equilibrio ecologico secondo la linea tenacemente difesa da Aurelio Peccei, al progresso sociale e civile secondo la linea ispirata a Lelio Basso (membro, tra l'altro, del Tribunale Russell); linea cui vorrei che tutti mirassero per aver diritto a goderne quanto a ciascuno può ragionevolmente spettare. Uno per tutti e tutti per uno, senza eccessive differenze e rivalità tra individui o classi o nazioni; rivalità utili soltanto se mirano a migliorare ovunque il benessere collettivo anziché curarsi soltanto di quello egoisticamente (e miopemente) individuale o settoriale o classista" (de Finetti 2015).

---

[1] Rimandiamo per questo a Fulvia de Finetti, Luca Nicotra, *Bruno de Finetti, un matematico scomodo,* Salomone Belforte, LivoBruno de Finetti, Riflessioni sul futuro, Civiltà delle macchine, anno XVI, 1968, n. 3, p.82. Riprodotto in G. Amari, F. de Finetti (rno 2009*,* e alla inclusa vasta bibliografia. Per la bibliografia si veda anche quella riportata in Bruno de Finetti, *L'invenzione della verità,* Raffaello Cortina Editore, Milano 2006. Che segue la suddivisione proposta dallo stesso de Finetti per l'opera *Raccolta degli scritti di Bruno de Finetti* (INA, Roma 1979, 10 volumi più volume di indici). Si contano circa 350 scritti che trattano di: teoria e calcolo della probabilità; matematica attuariale; matematica; economia e ricerca operativa; automazione; didattica e argomenti vari. Sono elencati inoltre 20 volumi di trattati, testi e monografie. Da consultare ovviamente il sito www.brunodefinetti.it curato da Fulvia de Finetti.

E ancora dice di sé: "Egli vorrebbe... che ognuno fosse lieto di poter ripetere tra sé: ' Io ho quel che ho donato ', e fosse commosso di poter dire, simmetricamente, ' Io sono grato per ciò che mi è stato regalato' " (de Finetti 1981).

2. Matematica e probabilità: pilastri del suo impegno in campo economico

Nella mirabile prolusione, tenuta nell'anno accademico 1948-1949 all'Università degli Studi di Trieste, che riscosse al tempo larga eco, dal titolo significativo "La funzione vivificatrice della matematica", il Nostro invitava ad abbandonare la visione "predicativa" della matematica per assumere quella "dialettica":

"[...] non più cioè una dottrina che si svolge partendo da un fondamento garantito una volta per sempre, ma una dottrina che, ad ogni istante del suo sviluppo storico, appare rispondente a un complesso di esigenze intellettuali ed applicative che può contenere nella sua evoluzione i germi di future crisi. Il punto di vista storicistico, di cui con tanta dottrina e tanto calore l'Enriques lumeggiò l'importanza per la comprensione dello sviluppo della matematica, s'insinuerebbe in tal modo fin dalla sua base gnoseologica" - e citando il Kasner - ' La matematica non è più riguardata come la chiave della Verità con la V maiuscola; essa può essere ritenuta come un piccolo, incompleto, ma enormemente utile Baedeker in una terra ignorata' "[2] (Annuario Università degli Studi di Trieste 1949).

Dopo aver ricordato la similitudine di Whitehead che paragona il ruolo della matematica nella storia del pensiero umano a quello di Ofelia nell'Amleto, un ruolo pressoché essenziale nell'opera teatrale, molto seducente e un poco folle, BdF concludeva: "Ma è, nel nostro caso, la follia apparente di chi costruisce per l'avvenire precorrendo ciò che ad altri non è ancora dato intravvedere. La matematica sgretola, scava, corrode con la sua critica le certezze di oggi, il cui crollo ci può atterrire, ma essa sta già sempre tessendo, spesso anche senza rendersi conto di tale destinazione la

[2] Discorso inaugurale per l'anno accademico 1948-1949 letto nell'Aula Magna dell'Università degli Studi di Trieste il 5 dicembre 1948. *Annuario dell'Università degli Studi di Trieste*, 1949, 19-34. Riprodotto in La matematica nella società e nella cultura. Rivista dell'Unione Matematica Italiana, cit., p. 282.

tela di ragno della nuova provvisoria certezza"(Annuario Università degli Studi di Trieste 1949).

Tra gli esempi che BdF citava nella sua prolusione in merito alla funzione disgregatrice ed insieme creativa della matematica, menzionava l'irruzione dell'incertezza e della probabilità nel pensiero moderno e non solo scientifico.

"L'incertezza domina ovunque. Tutta la nostra vita è immersa nell'incertezza; nulla – all'infuori di ciò – si può affermare con certezza"(de Finetti, Emanelli 1967)[3].

E ci propone la sua concezione soggettiva della probabilità come il miglior modo per fronteggiarla consapevolmente.

Concezione, la cui espressione più compiuta ed organica si trova nei due volumi del suo Teoria delle probabilità, considerato da Lindley, nella prefazione alla edizione inglese,"*one of the great books of the world*"[4] (de Finetti 1970).

---

[3] Bruno de Finetti, L'economia dell'incertezza, Prima parte del volume Bruno de Finetti, Filippo Emanuelli, *Economia delle assicurazioni*, Utet, Torino 960, p. 3. Si tratta di un saggio magistrale che, partendo dal riconoscimento della pervasiva incertezza nella vita e dal concetto della probabilità soggettiva, affronta, con la sua ben nota coerenza, le tematiche delle decisioni, dell'informazione, del rischio, del differimento, dell'attualizzazione, per arrivare a quella più specifica dell'assicurazione privata; da tener ben distinta dalla "sicurezza sociale" (preferita al termine "assicurazione sociale") che, proprio perché affronta temi in una dimensione dell'economia come un tutto e di norma con il vincolo dell' "obbligatorietà", come quello previdenziale, non risponde e non può logicamente rispondere a meri criteri di gestione aziendale. Sicurezza sociale, da considerare inoltre come "strumento essenziale" e come "fattore più universale e politicamente basilare" di un "programma antialeatorio" che fronteggi l'incertezza e le sue perniciose conseguenze, ivi comprese le diseguaglianze.

[4]Lo ricorda nella sua recensione ai volumi I.J. Good che conclude : "In summary, these volumes make important writings of this pioneer available to the English-reading world, and will encourage some probabilists, statisticians, and philosophers of science to learn Italian". Come fece d'altronde il suo amico Savage. I.J. Goods, Bulletin of the American Mathematical Society, volume 83, Number I, January 1977. La casa editrice Wiley, a Marzo 2017, ha ristampato in volume unico *Theory of Probability*, disponibile anche in versione E.book.

Secondo Marco Mondadori, nella prefazione ad una importante raccolta di scritti di BdF, da lui curata, si tratta del "nucleo di un programma di ricerca scientifica tra i più progressisti del '900 che ha ormai completamente rovesciato quel pilastro di 'conoscenza stabilità che era la credenza superstiziosa nella probabilità oggettiva'[5] (de Finetti 1989).

Programma di ricerca che comincia appunto nella liberazione da "una delle credenze superstiziose più pericolose ed aberranti, quella che ammette e afferma che esistano delle 'probabilità oggettive', e cioè che la probabilità misuri qualche proprietà oggettiva che sta fuori di noi, nel mondo esterno.

"La probabilità va vista piuttosto come il grado di fiducia di un dato soggetto, in un dato istante e con un dato insieme di informazioni, riguardo al verificarsi di un dato evento"(de Finetti 1989)..

"La novità essenziale nel metodo scientifico – afferma BdF - sarebbe la sostituzione della logica con il calcolo della probabilità; al posto di una scienza razionalistica in cui si deduce il certo dal certo si avrebbe una scienza probabilistica in cui si deduce il probabile dal probabile"(de Finetti, Nicotra 2008).

Si comprende quindi meglio la chiusura del suo giovanile ma importante lavoro filosofico e scientifico:

"Si vede che tutto è costruito sulle sabbie mobili, benché naturalmente si cerchi di poggiare i pilastri sui punti relativamente meno pericolosi;

---

[5] Marco Mondadori, Prefazione a *Bruno de Finetti, La Logica dell'incerto*, Il Saggiatore, Milano 1989. Il volume raccoglie importanti scritti di BdF sulla probabilità. Con l'analogia di questa con la "scomessa equa", e con il suo "teorema della rappresentazione", usando il concetto di "scambiabilità", ricondusse le valutazioni soggettive a logica coerenza, evitando i rischi dell'arbitrarietà e i limiti di applicabilità soprattutto nei casi di eventi non ripetibili. Superando, nel contempo, aporie e contraddizioni delle concezioni classica e frequentista, in una visione più generale anche nei confronti di altre impostazioni soggettivistiche, come quella cambridgeana. Si veda a tal proposito, nel medesimo volume, lo scritto Probabilisti di Cambridge (pp. 203-222). Con particolare riguardo ai rapporti con l'economia, si veda, oltre a quanto riportato in nota 7, Bruno de Finetti*, Un matematico e l'economia*, Franco Angeli Editore, Milano 1969. Importante, di Giorgio Lunghini, "Bruno de Finetti e la teoria economica". Relazione presentata alla Accademia dei Lincei nella *Bruno de Finetti Centenary Conference*, 15-17 novembre 2006. Indirizzato alle scuole medie superiori, Bruno de Finetti, Ferruccio Minisola, *La matematica per le applicazione economiche*, Edizioni Cremonese, Roma, 1961.

comunque, non soltanto le leggi e le previsioni non sono certe, ma solo probabili, ma anche il fatto che certi schemi in cui riteniamo opportuno rappresentare i fenomeni, come gli stessi concetti di spazio e tempo, e i criteri di misura delle distanze e dei tempi, continuino effettivamente a mostrarsi opportuni o anche soltanto conservabili, è un fatto che non può considerarsi certo, ma soltanto probabile (sia pure immensamente probabile)"[6](de Finetti 2006).

È una concezione che si pone *naturaliter* all'intersezione delle tre dimensioni filosofiche della logica, dell'etica e dell'estetica[7] (de Finetti, Dorfles, Nervi 1966).

Ancor più e più in generale che per i teoremi di Gödel, a cui pur allude nella richiamata prolusione insieme al dibattito sulla crisi dei fondamenti della matematica, la concezione epistemologica di de Finetti postula il concorso di altre discipline, aborrendo pretese certezze, ingiustificate autoconsistenze e pericolosi imperialismi di disciplina.

Senza poter ricordare i nomi dei tanti studiosi di varie discipline coinvolti nei corsi CIME di cui si parlerà più oltre, ci piace menzionare l'attenzione da parte di un cognitivista e filosofo della scienza come Massimo Piattelli Palmarini, e di un autorevole fisico come Carlo Rovelli[8] (Piattelli Palmarini 1987).

---

[6] Bruno de Finetti, *L'invenzione della verità*, cit., p. 146. Il volume, con importante introduzione di Giordano Bruno e Giulio Giorello, pubblica un saggio inedito del 1934 del giovane BdF, recuperato grazie a Fulvia de Finetti, che illustra i consolidati convincimenti filosofici ed epistemologici alla base dei suoi contributi scientifici.

[7] In merito più specificatamente alle questioni estetiche, si veda l'acuto e stimolante dialogo tra BdF, Gillo Dorfles e Pier Luigi Nervi, in Forme estetiche e le leggi scientifiche, Civiltà delle macchine, n. 14, 3 1966. Nell'esprimere contrarietà "alla costruzione di teorie estetiche, specie se basate su chiacchiere filosofiche o pseudofilosofiche", BdF delinea una visione dell'estetica (e dell'arte) come replica della fantasia creatrice alle (anzi grazie alle) ristrettezze fisiche e culturali del tempo e dello spazio ad essa contemporanee. Per maggiori argomentazioni sia permesso il riferimento a G. Amari, Unità, coerenza e fecondità del pensiero di BdF, sul numero n. 1, 2017 di Economia & Lavoro, in corso di pubblicazione.

[8] Massimo Piattelli Palmarini, *S come cultura, Protagonisti, luoghi e idee delle scienze contemporanee*, Mondadori, Milano, 1987, Bruno de Finetti, il soggetto

Questa condizione ontologica, permeata di incertezza, in un campo continuo senza soluzioni o peggio dicotomie[9], avvalora in pieno la funzione epistemologica della statistica bayesiana che, come è noto, ci aiuta a procedere nel tempo, modificando le nostre stime probabilitiche alla luce di maggiori informazioni, esperienze e confronto aperto con l'altro; non per raggiungere impossibili verità oggettive, ma solo provvisori convincimenti intersoggettivi. Possiamo, anzi dobbiamo tentare di essere "obiettivi", ma non possiamo essere "oggettivi" (de Finetti 1962).

Il che ci riporta a quella concezione "storicistica" sostenuta all'inizio, consapevoli cioè del "farsi della storia", per citare un concetto gramsciano, che non ha nulla a che vedere con leggi deterministiche dal destino segnato sia in termini catastrofici che in quelli del progresso inarrestabile.

Ce lo ricorda lo stesso de Finetti quando avverte che "la valutazione di probabilità, non solo si deve riferire ad *un* evento (in senso *specifico*), ma dipende anche dall'insieme variabile di circostanze ritenute rilevanti rispetto al suo verificarsi, note al momento (e, in genere, varianti di momento in momento)". E se usualmente non si evidenzia tale subordinazione di carattere generale, perché scontata, non può mai essere dimenticata se si vuole avere piena contezza della vera e provvisoria condizionalità delle nostre stime[10] (de Finetti 1989).

Va abbandonata dunque l'etica della Verità, ma abbracciata quella del dubbio, della sincerità, della lealtà [11] (Zagrebelsky 2008), nel coraggio e nel dovere dell'opinione informata e dichiarata.

---

della probabilità, pp. 147-151. Carlo Rovelli, Sì, no, anzi: probabilmente. Il Sole 24 Ore, Domenica 20 gennaio 2013. Sempre di Rovelli, L'incertezza per compagna di viaggio, La Lettura del Corriere della Sera, 6 novembre 2016.

[9] Come spesso ancora avviene pensando di distinguere nettamente il "rischio" dall' " incertezza"; dove il primo si può trattare con metodi stocastici, usati non di rado con ingenua ed improvvida sicurezza, mentre per la seconda nulla si potrebbe dire o fare.

[10] Bruno de Finetti, La probabilità: guardarsi dalle contraffazioni!, in part. Paragrafo 7, Probabilità: suo carattere relativo. Testo dell' "ultima lezione" tenuta in occasione del collocamento "fuori ruolo" all'Istituto Matematico G. Castelnuovo il 29 novembre 1976. Pubblicato in Scientia, 111, 1976, pp. 255-281. Riprodotto in *La logica dell'incerto*, cit., pp. 149-188.

[11] Di un'etica del dubbio parla anche Gustavo Zagrebelsky, *Contro l'etica della verità*, Laterza, Bari 2008. Un'etica del dubbio che è per una verità che deve essere sempre riscoperta, e quindi mai assoluta e definitiva.

Massimo De Felice sottolinea: "[…] la dimensione 'etica' del suo probabilismo, *l'impegno alla chiarezza (contro i 'vari modi di non dire nulla'), il non sottrarsi ad avere e dichiarare e sottoscrivere l'opinione, l'obbligo della coerenza, il superamento delle pigrizia (cercando l'altra via oltre quella che per prima ci viene in mente)*" (de Felice 2019).

Il dovere all'*opinione* esercitato sui nostri stessi presupposti di valore - da considerare in modo non assoluto e sempre soggetti a conferma o revisione - e sulle conseguenze future delle nostre azioni, è anche il modo più serio di confrontare l'etica dei principi con l'etica della responsabilità.

Un impegno a cui si attennero intellettuali come BdF che adoperarono il coraggioso linguaggio della verità *come la vedevano*, nella luce crepuscolare della probabilità, pagando spesso incomprensioni ed isolamento[12] (Massarenti 2011). Quelli che Ralph Dahrendorf chiamerà Erasmiani (Dahrendorf 2007).

In questa luce vanno anche viste molte delle sue inizative sul piano scientifico, economico e civile.

Ricordiamo innanzitutto il lungo ciclo decennale dei corsi CIME (de Finetti, Amari 2015), che rappresentarono un avvenimento scientifico e culturale di grande rilievo, non solo nazionale. Ad essi parteciparono autorevoli economisti stranieri, compreso R. Frish, insignito con il primo premio Nobel per l'economia nel 1969 insieme a J. Tinbergen; e molti giovani studiosi italiani che rappresenteranno poi buona parte del mondo accademico italiano.

Svolti sotto la direzione esclusiva di BdF, per conto del CIME (Centro Internazionale Matematico Estivo) e finanziati dal 1966 dall'Ente per gli studi monetari, bancari e finanziari "Luigi Einaudi", diretto in quegli anni da Federico Caffè che ne fu anche partecipe collaboratore, quei corsi, diventati da allora convegni CIME-EINAUDI, rappresentarono un vero e proprio progetto di ricerca sociale squisitamente definettiano; sempre più finalizzato, con impostazione multidisciplinare, "fusionista" come la chiamava,

[12] "Nel crepuscolo della probabilità", è il significativo titolo di un saggio di Giulio Giorello dedicato a BdF, in A Massarenti (a cura), *Qualcosa di grandioso, l'infinita bellezza e complessità di tutto ciò che esiste,* Dalai editore, Milano, 2011, pp. 185-209.

all'avanzamento civile e al miglioramento della condizione umana[13](de Finetti 1959).

Gli stessi titoli dei convegni e le relazioni di BdF sono significativi al riguardo.

"Requisiti per un sistema economico accettabile in relazione alle esigenze della collettività", è il titolo del VII convegno tenuto ad Urbino dal 20 al 25 settembre 1971; e "L'Utopia come presupposto necessario per ogni impostazione significativa della scienza economica", è il titolo della relazione del direttore, in perfetta sintonia con lo spirito del programma (de Finetti 1973).

Ma nel volume che pubblica nel 1973 gli atti, si propone anche un aggiunta del maggio 1972 al titolo dell'anno precedente, "come spunto per riprendere le discussioni in argomento nel prossimo convegno" che si preannunciava per il 3-8 luglio 1972: " … E per ogni tentativo di salvare l'umanità dall'autodistruzione ".

In questa aggiunta BdF fa largo richiamo al noto Rapporto del MIT sui "Limiti dello sviluppo", ma con un'impostazione in parte diversa dando "alle riforme sociali un valore prioritario per ragioni anzitutto morali … e non tanto per evitare la rovina". Si condivideva comunque la richiesta del Rapporto per una "non cieca opposizione al progresso, bensì ad un cieco progresso". Dandone una lettura non stagnazionista o peggio di decrescita. E ne riporta la conclusione : "[…] Appena sarà sviluppata una più penetrante conoscenza delle condizioni e dei modi di funzionamento del sistema del mondo, queste nazioni [sviluppate] si renderanno conto che, in un mondo che ha essenzialmente bisogno di stabilità, *i loro alti livelli di sviluppo possono venire giustificati e tollerati soltanto se essi servono non come trampolini di lancio per raggiungere altri ancor più in alto, bensì come piattaforme dalle quali organizzare una più equa distribuzione della ricchezza e del reddito su tutto il mondo"*. E BdF da parte sua commenta: "Mi sembra il vero ed unico modo di ripresa e di traduzione in un piano

---

[13] Una programmata impostazione "fusionista" nell'insegnamento della matematica si può trovare nella sua importante Prefazione alla prima edizione del volume universitario *Matematica logico intuitiva*, Editrice scientifica Triestina, Trieste, 1944 (seconda edizione, Edizione Cremonese, Roma 1957; ristampa Giuffrè Editore, Milano 2005). E nella Prefazione e capitolo primo del testo per le scuole medie superiori, Bruno de Finetti, Ferruccio Minisola, *La matematica per le applicazione economiche*, cit.

concreto ed organicamente completo, dell'alto indimenticabile messaggio della *Pacem in Terris*, l'unica proposta idonea per promuovere e sperabilmente costruire un'effettiva '*pacem in terris*' ".

"Crisi dell'energia e crisi di miopia", è il titolo del X Convegno CIME – EINAUDI tenuto ad Urbino dal 2 al 7 settembre 1974. Nella sua relazione BdF precisa che "I problemi del futuro implicano il superamento dei criteri di interesse gretto e immediato, in favore di visioni più lungimiranti nel tempo e nello spazio, visioni riguardanti il bene comune per tutti gli esseri viventi (non solo uomini, ma tutti gli animali e vegetali) su tutta la terra e per tutto il tempo futuro in cui la vita potrà essere conservata".

E per un'esatta comprensione di tale visione riteneva necessario considerare anche l'aspetto religioso: "di religione non come mistificazione formalistica-rituale-superstiziosa, ma come senso di responsabilità indirizzato al bene comune secondo il Cantico delle Creature di Santo Francesco".

"Dall'utopia alla alternativa (1971-76)", è il titolo dell' XI Convegno CIME – EINAUDI tenuto a Urbino dal 1 al 6 settembre 1975 (de Finetti 1976)**.**

In esso, BdF continua ed approfondisce il suo programma, con radicalità e impazienza più accentuata rispetto al convegno precedente, riallacciandosi in continuità con quello del 1971, come si legge dalle stesse date indicate nel titolo. È ormai sempre più urgente che dall'Utopia si passi all'Alternativa.

Si legge, nell'ultima di copertina del volume che pubblica gli atti del convegno, un testo scritto evidentemente da BdF, che dà piena contezza dello spirito che anima il progetto definettiano:

"Il cammino 'Dall'utopia all'alternativa', indicato nel titolo, si riferisce allo sviluppo di tematiche e discussioni avvenuto dal 1965 al 1975 (e continuerà, per lo meno nel 1976[14](de Finetti 1976)) nei Corsi (poi Convegni) di economia matematica organizzati dal CIME (Centro Internazionale Matematico Estivo). Esaminare, e cercare come si potrebbe

[14] Sembra ci sia stato ancora un "Convegno ex Urbino", il 15/16 settembre 1977 a Roma (per riduzione di finanziamenti), come si legge da una traccia di relazione ("Parole introduttive") di BdF. In questa traccia BdF riprende ed accentua le ben note tematiche, arrivando ad auspicare una convergenza delle forze della sinistra per quei mutamenti sociali che gli stavano a cuore.

risanare, il disfunzionamento dell'economia, discutere le pretesemente indiscutibili e sacre 'leggi economiche' e istituzioni economiche, appariva, inizialmente, un'eresia, comunque un'utopia, cioè una visione lontana nell'avvenire (seppure mai realizzabile, come avrebbero sentenziato i soliti sedicenti 'benpensanti')[15] (de Finetti, Amari 2015).

Ma nella recente enciclica *Evangelium Gaudium*, si legge "I cittadini vivono in tensione tra la congiuntura del momento e la luce del tempo, dell'orizzonte più grande, dell'utopia che ci apre al futuro come causa finale che ci attrae. Da qui emerge un primo principio per progredire nella costruzione di un popolo: il tempo è superiore allo spazio"(Papa Francesco 2013). E non minore sintonia avvertirebbe con l'ultima enciclica di dichiarato spirito francescano, *Laudato sì* (Papa Francesco 2015), lo stesso a cui BdF, come abbiamo visto, faceva esplicito riferimento.

L'antropologia considerata da BdF non è quella dello "sciocco razionale" (Sen 1986), ma quella che si trova nel capolavoro del Cervantes, di Don Chisciotte con la sua generosa fantasia utopica e di Sancho Panza dalla terrena quotidianità. Ma l'uno indispensabile all'altro, tanto è vero che al rinsavimento del primo (che quindi muore), fa riscontro l'alienazione del secondo. Claudio Magris, in un suo magistrale saggio, ce ne dà questa profonda interpretazione esistenziale e ci invita a non rinunciare all'utopia e alle ideologie, nonostante e malgrado il "disincanto", conseguenza di aspettative deluse e di tragici errori commessi (Magris 1999).

La filosofa Agnes Heller, con Riccardo Mazzeo, passa in rassegna le tante utopie e distopie antiche e recenti con le annesse speranze e paure, e invita comunque a non abbandonare la speranza nell'immaginazione sociale. Ricorda che al "vento" positivo delle utopie è spesso seguito il "vortice" delle "distopie". Ma da esse dobbiamo imparare, "[...] non disperare, non cedere, senza però inseguire neppure vuote ilusioni... Utopia [da intendere oggi] come coraggio civico. Lo slogan 'coltivate il vostro giardino' non suggerisce l'abbandono dell'immaginazione socio-storica, ma la responsabilità assunta per l'immaginazione sociale. Si può essere responsabili dell'immaginazione sociale *se l'immaginazione si basa sulla*

[15] Questo programma lo voleva estendere a livello internazionale, e ne scrisse al suo amico Franco Modigliani che gli suggerì i nomi di Leontieff, Vanek, Lundberg, Arrow. Si veda G. Amari, F. de Finetti, *Bruno de Finetti, un matematico tra Utopia e riformismo*, cit., p. 312.

probabilità [corsivo nostro] e non soltanto sulla possibilità o ancor meno sull'impossibilità" (Heller, Mazzeo 2016).

La coraggiosa immaginazione sociale dell'Utopia definettiana, che non scarta a priori disegni incautamente ritenuti impossibili, è basata sulla probabilità e ci immunizza dalle paurose distopie o dalle "vuote illusioni". Dal libero e aperto disegno dell'Utopia si passa, infatti, al coraggio del concreto e rigoroso riformismo. L'uso del paradigma paretiano ne è un esempio significativo[16] (de Finetti 1973).

3. Per un nuovo umanesimo

La pervasiva globalizzazione finanziaria richiede l'inaridimento culturale con la scomparsa o emarginazione di visioni critiche. Lo denuncia allarmata la filosofa Martha Nussbaum quando lamenta la progressiva mortificazione delle discipline umanistiche.

Paventa il rischio di avere perfetti ingegneri e provetti tecnici senza maturazione critica e tali da assumere come dato non discutibile il quadro economico ed istituzionale presente; alieni cioè da un'adeguata consapevolezza civile. Naturalmente, è ben lungi da sottovalutare l'importanza dell'educazione scientifica, ma questa deve essere inserita sempre in un contesto storico e rivendica che in tale insegnamento vadano sempre recuperati "gli aspetti umanistici".

---

[16] BdF riafferma il valore degli "ottimi" paretiani, manifestando però la sua sfiducia nel fatto che si possano raggiungere con il libero mercato. Inoltre, poiché il paradigma paretiano, rifiutando di esprimere giudizi di valore, prende per data ogni distribuzione di reddito e di ricchezza anche se fortemente ingiusta, ci possono essere ottimi in senso "tecnico", ma che possono essere considerati "non buoni". Si tratta, al contrario, di partire con obiettivi ritenuti accettabili e condivisi democraticamente, per poi passare a forme di ottimizzazione e di efficienza produttiva. Il tema è diffusamente trattato in *Requisiti per un sistema economico accettabile in relazione alle esigenze della collettività: Urbino 20-25 settembre 1971*, cit. Si veda anche il sintetico e lucido contributo del Nostro all'importante convegno internazionale su Vilfredo Pareto tenuto all'Accademia de Lincei a Roma il 25-27 ottobre 1973. *Atti dei Convegni Lincei*. Accademia Nazionale dei Lincei, Roma 1975, pp. 220-221.

"In realtà, anche quelli che potremmo definire come gli aspetti umanistici della scienza e della scienza sociale – l'aspetto creativo, inventivo, e quello di pensiero critico, rigoroso – stanno perdendo terreno, dal momento che i governi preferiscono inseguire il profitto a breve termine garantito dai saperi tecnico-scientifici i più idonei a tale scopo" (Nussbaum 2011).

"L'ideale – sosteneva Federico Caffè, grande amico di BdF - [è quello] di costruire un mondo in cui il progresso sociale e civile non rappresenti un sottoprodotto dello sviluppo economico, ma un obiettivo coscientemente perseguito" (Caffè 1976).

Da parte sua BdF ben argomentava:

L'Utopia […] consiste semplicemente, nel dire che si deve pensare dapprima ai *fini*, e precisamente ai fini veri ed ultimi, e solo *dopo* ai mezzi per raggiungerli. I fini veri ed ultimi non possono essere altro se non le effettive condizioni di vita per ogni singolo individuo e pertanto per la collettività. E *non*, come spesso viene suggerito, certe entità macroeconomiche o forme di istituzioni e di strutture scambiate per obiettivi anziché tutt'al più, come strumenti più o meno validi per descrivere sommariamente e per conseguire l'obiettivo reale. Va menzionato – come esempio estremo di meschina ottusità – la pur diffusa consuetudine (anche nelle riverite sfere ufficiali statali e superstatali) di esprimere gli obiettivi in termini di 'GNP' (prodotto nazionale lordo). Che è veramente 'lordo', non solo nel senso mercantile del termine, ma anche in quello morale, dato che vi si assommano indiscriminatamente tutte le cose che hanno un prezzo in quanto 'utili', vuoi come alimenti per i bambini o come armi per le imprese dei gangster o per qualunque altro fine: se (e soltanto se) qualcuno che ha soldi è disposto a pagarla, una cosa è un *bene* 'economico' "[17](de Finetti 1976).

E ancora:

"Qualcuno potrebbe sintetizzare tale ideale parlando di un *diritto* assicurato a tutti di quanto occorre per vivere decentemente, come corrispettivo al *dovere* di prestare la propria opera al servizio della collettività al meglio delle possibilità di ciascuno […]. Chi godesse solo dei 'benefici' senza corrispondere con prestazioni date volonterosamente e

[17] Bruno de Finetti, *Dall'Utopia all'Alternativa*, cit. pp. 11-12. Troviamo anche qui idee originali e precorritrici sulla corretta individuazione dei mezzi e dei veri fini, sul rapporto tra microeconomia e macroeconomia, sul significato e le contraddizioni del PIL.

lietamente (anche quando costano sacrificio), o chi dovesse subire il lavoro o una mansione o un sacrificio come costrizione alienante, senza il dovuto riconoscimento materiale e morale, non vivrebbe in un mondo sufficientemente riscattato dalle vergogne delle epoche passate, lontane e meno lontane (schiavitù, corvées, condizioni di lavoro inumane, colonialismo), o presenti (insicurezza del lavoro, lavoro alienante, inesistenza o forme insufficienti o male organizzate di assistenza, ecc.)**"** (de Finetti 1976).

Riducendo al "minimo, drasticamente, l'ambito delle cose in cui si applica la funzione del denaro impedendogli di dar luogo – come ora con speculazioni affarismi e manovre – a far funzionare il 'mulino del diavolo' " de Finetti 1976).

4. Conclusione

A tale avanzamento civile e culturale sono destinati dunque la matematica di BdF e il suo impegno ad insegnarla perché: "... farla comprendere significa anzitutto farla amare, farla sentire non avulsa dai pensieri e meditazioni e preoccupazioni d'ogni giorno, ma ad essi siffattamente frammista da far apparire all'opposto arido e opaco il pensiero che non sappia attingere alla sua luce"[18] (de Finetti 1959).

Così, al "vento" dell'utopia definettiana, sfuggendo ai "vortici" delle distopie, alla luce degli ideali e nel "crepuscolo della probabilità", possiamo dire con il poeta:

[18] Dalla Prefazione a *Matematica logico intuitiva*, cit., p. X. Paradigmatico del suo modo di intendere la didattica è *Saper vedere in matematica*, Loesher, Torino 1967 (1986). Ripubblicato in La Matematica nella società e nella cultura. Rivista dell'Unione Matematica Italiana. Serie I, Vol. VIII, Dicembre 2015, cit., che raccoglie scritti di BdF e di altri autori sulla didattica definettiana. Organizzate da Proteo Fare Sapere e dalla Fondazione Giuseppe Di Vittorio, con l'attiva partecipazione della prof.ssa Carla Rossi e del prof. Alberto Zuliani, ci sono state, il 12-13 ottobre 2017 a Firenze, due giornate dedicate all'insegnamento di BdF. Venne rappresentata, in tale occasione, la bella pièce teatrale "Probabilmente... de Finetti", prodotta, scritta e recitata dalla compagnia L'Aquila Signorina di Gabriele Argazzi e Barbara Bonora.

passa una vela,
spingendo
la notte
più in là[19].

Bibliografia

Amari, G., de Finetti, F. (2015), Bruno de Finetti, un matematico tra Utopia e Riformismo, Ediesse, Roma.

Annuario dell’Università degli Studi di Trieste (1949), 19-34, p. 292.

Caffè, F. (1976), *Problemi controversi sull’intervento pubblico nell’economia*, Note economiche, n. 6.

Dahrendorf, R. (2007), *Erasmiani, gli intellettuali alla prova del totalitarismo*, Laterza, Bari-Roma.

de Felice, M., Giorello, G., Moriconi, F., Piccinato, L., Salinetti, G. (2019), *Conoscere de Finetti, per il governo dell’incertezza*, Mondadori Università Sapienza, Milano.

de Finetti, B. (1959), *Matematica logico intuitiva*, Cremonese.

de Finetti, B. (1962), *Obbiettività e oggettività, critica ad un miraggio*, La Rivista trimestrale, anno I, n.2.

de Finetti, B. (1967), *Saper vedere in matematica, Loesher*, Torino.

de Finetti, B. (1969), *Un matematico e l’economia*, Franco Angeli Editore, Milano.

[19] La poesia è di Tonino Milite.

de Finetti, B. (1970), *Teoria delle probabilità*, Einaudi, Torino, p. XV.

de Finetti, B. (1973), *Requisiti per un sistema economico accettabile in relazione alle esigenze della collettività: Urbino 20-25 settembre 1971*, Franco Angeli Editore, Milano.

de Finetti, B. (1975), *Crisi dell'energia e crisi di miopia*, Franco Angeli Editore, Milano 1975.

de Finetti, B. (1976), *Dall'utopia all'alternativa (1971-76)*, Franco Angeli Editore, Milano.

de Finetti, B. (1981), *Scritti (1926-30)*, Cedam, Padova, pp. XVII-XXIV.

de Finetti, B. (1989), *La Logica dell'incerto*, Il Saggiatore, Milano

de Finetti, B. (2006), *L'invenzione della verità*, Cortina Raffaello, p. 146

de Finetti, B. (2015), Chi sono io ?, *La Matematica nella società e nella cultura*, Rivista dell'Unione Matematica Italiana. Serie I, Vol. VIII, 463-467.
de Finetti, B., Emanuelli, F. (1967), Economia delle assicurazioni, Utet, Torino, p. 3.

de Finetti, B., Dorfles, G., Nervi, P.L. (1966), *Forme estetiche e le leggi scientifiche*, Civiltà delle macchine, n. 14, 3.

de Finetti, B., Minisola, F. (1961), *La matematica per le applicazione economiche*, Edizioni Cremonese, Roma.

de Finetti F., Nicotra L. (2008), *Bruno de Finetti Un matematico scomodo*, Belforte, Livorno, p.135

Goods, I.J. (1977), *Bulletin of the American Mathematical Society*, volume 83, Number I

Heller, A., Mazzeo, R. (2016), I*l vento e il vortice, Utopie, distopie, storia e limiti dell'immaginazione*, Edizioni Centro Studi Erickson S.p.A., Trento, pp. 76-77.

Magris, C. (1999), *Utopia e disincanto*, Saggi 1974-1998, Garzanti, Milano, pp. 7-16.

Massarenti, A. (2011), *Qualcosa di grandioso, l'infinita bellezza e complessità di tutto ciò che esiste*, Dalai editore, Milano, pp. 185-209.

Nussbaum, M. (2011), *Non per profitto. Perché le democrazie hanno bisogno della cultura umanistica*, Il Mulino, Bologna, p. 22.

Papa Francesco (2013), *Evangelii Gaudium*, Libreria Editrice Vaticana, Città del Vaticano, p. 229.

Papa Francesco (2015), *Laudato sì, Enciclica sulla cura della casa comune*, Libreria Editrice Vaticana, Città del Vaticano.

Piattelli Palmarini, M. (1987), *S come cultura, Protagonisti, luoghi e idee delle scienze contemporanee*, Mondadori, Milano.

Sen, A.K. (1986), *Sciocchi razionali: una critica dei fondamenti comportamentistici della teoria economica*, in Scelta, benessere, equità, Il Mulino.

Stiglitz, J., Sen, A.K., Fitoussi, J.P. (2010), *La misura sbagliata delle nostre vite Perché il PIL non sufficiente per valutare benessere e progresso sociale: Rapporto della commissione per la misurazione della performance economica e del progresso sociale*, Etas Milano.

Zagrebelsky, G. (2008), *Contro l'etica della verità*, Laterza, Bari.

G. & L. E. R., Vol. 22 No. 2, 2018

Flavio Pressacco* - Paolo Serafini† - Laura Ziani‡

DE FINETTI AND MARKOWITZ MEAN VARIANCE APPROACH TO REINSURANCE AND PORTFOLIO SELECTION PROBLEMS: A COMPARISON

Abstract

Based on a critical analysis of de Finetti's paper, where the mean variance approach in finance was early introduced to deal with a reinsurance problem, we offer an alternative interpretative key of such an approach to the standard portfolio selection one. We discuss analogies and differences between de Finetti's and Markowitz's geometrical approaches.

JEL CLASSIFICATION: B23; C61; D81; G11; G22;

KEYWORDS: Finance; Mean-variance Efficiency; Constrained Quadratic Optimization; Portfolio Selection; Proportional Reinsurance.

* Dept. of Economics and Statistics, Udine University, via Tomadini, n. 30/a, Udine, Italy. Phone: +39-0432-249310. *E-mail address*: flavio.pressacco@uniud.it.

† Dept. of Mathematics and Informatics, Udine University, via delle Scienze, n. 8, Udine, Italy. Phone: +39-0432-249307. *E-mail address*: laura.ziani@uniud.it.

‡ Dept. of Economics and Statistics, Udine University, via Tomadini, n. 30/a, Udine, Italy. Phone: +39-0432-249307. *E-mail address*: laura.ziani@uniud.it.

1. Introduction

H. Markowitz was awarded the Nobel Prize in Economics thanks to his three papers (Markowitz 1952, 1956, 1959) in the 1950s in which he introduced the mean-variance approach in quantitative finance in order to solve a problem of a portfolio of risky assets' selection (henceforth investment problem).

Only recently it has been discovered (see Rubinstein, 2006a) that the primacy of the mean-variance approach in finance should be credited to B. de Finetti, who introduced it in a paper (de Finetti, 1940) concerning the selection of individual proportional reinsurance of a set of risks in an insurance company portfolio (henceforth reinsurance problem).

While obviously Markowitz's papers are very well known in financial world, the one of de Finetti, written in Italian language in an actuarial journal at the beginning of the Second World War, went unnoticed to researchers in financial economics and its knowledge remained restricted for a long time to European actuarial circles; this explains its late rediscovery[1].

At first sight, the investment and the reinsurance problems seem to be quite different; on the contrary, they reveal to have much in common once they are expressed in a formalized version. Hence, one of the goals of this paper is an analysis of analogies and differences between the authors' approaches. A lot of analogies may be found mainly in the first paper (Markowitz, 1952), which is largely based on geometric intuitions, a method often privileged by de Finetti too, at least in his papers devoted to economic and financial applications.

It is convenient to recall here that while Markowitz did not stop investigating the topic, thus being able to apply also newly developed optimization techniques (Karush, 1939; Kuhn-Tucker, 1951), de Finetti completely forgot his paper[2], so that he did not even consider either to purge

[1] An early discussion of the connections between de Finetti's work and portfolio theory can be found in Pressacco (1986) as quoted in Rubinstein (2006a). An English translation of the first part of de Finetti's paper has been provided in Barone (2006).

[2] He himself (de Finetti, 1969) did not include it in a list of his own papers bearing some, even minor, connection with economic topics.

it from shortcomings or to extend results to other fields of financial economics.

With this respect, another goal of the paper is to discuss how these shortcomings can be fixed and which enlightening interpretations can be also obtained in the investment problem through further investigation (but remaining faithful to de Finetti's logic).

The plan of the paper is as follows: Section 2 offers a preliminary largely informal glance to the analogies and differences in the authors approach to the problem. Section 3 is devoted to a quick recall of de Finetti's reward risk approach to the reinsurance problem, followed in section 4 by an informal resume of his friendly procedure to find the whole efficient set in the asset space as well as of some shortcomings, signalled by Markowitz (2006) about its applicability to a general correlation structure in the absence of a regularity condition. Section 5 shows the closed form formulae of the reinsurance problem for the case of no correlation and group correlation respectively. Section 6 shows how to adjust de Finetti's rules when the regularity hypothesis does not hold. In the next sections 7 (tools), 8 (regular case), 9 (non regular case), the extension of de Finetti's approach to the standard investment problem is presented in a quite informal way, while a more technical presentation is given in the Appendix. Finally, in section 10 a discussion of some examples offers an enlightening comparison of Markowitz and de Finetti's geometrical intuitions behind the critical line algorithm in the investment problem. Conclusions follow in section 11.

2. A preliminary glance at the problem

Let us start from a preliminary statement concerning the strict connection between the (re)insurance and the investment problem.

An insured risk with liability X_i in a given single period time horizon may be dealt with as a risky asset with random gain G_i, if we consider the difference $G_i = P_i - X_i$ where P_i is the premium received to insure the risk (net of loading expenses but gross of positive safety loadings, so that $P_i > E(X_i)$. A proportional reinsurance on original terms of a quota $1 - x_i$ of the risk, i.e. retention of a quota x_i, originates a random post-reinsurance gain $x_i G_i$, and an overall post-reinsurance gain $\sum_i x_i G_i$.

On the other side, a portfolio of risky assets can be defined from a set of quotas x_i of the disposable wealth of an investor. If the assets have a single period random rate of return G_i, the random rate of return on the portfolio is $\sum_i x_i G_i$. Standard constraints are $0 \le x_i \le 1$ for the reinsurance case and $x_i \ge 0$, $\sum_i x_i = 1$ for the investment one. These constraints were applied by de Finetti (de Finetti, 1940) to the reinsurance problem faced by an insurance company and in the Markowitz (1952) to the investment one, while in subsequent papers he considered also other constraints, such as industrial ones and so on (for a general treatment of affine constraints see Markowitz, 1987, chapter 6).

In this framework, a first strict analogy between the authors is the choice of an integrated reward-risk approach. Indeed both were unsatisfied with the one-sided approach to decisions, largely prevailing at that time, which was risk driven in the (re)insurance field (where decisions aimed to keep conveniently bounded the ruin probability of the insurance company), and reward driven in the investment sector (a consequence of the wrong idea that naive diversification could fully get rid of any risk[3]).

The second strong analogy concerns the reward and risk measures to be used in the analysis. Both authors reflected carefully on the problem and their common final choice was expectation and variance of the random gain (or rate of return) in a single period horizon, even if de Finetti came to the variance only as a (second best) computationally more convenient, equivalent substitute of the (first best) ruin probability risk measure, and Markowitz dedicated an entire chapter (chap. 9) of his book (Markowitz, 1959) to deal with a mean-semideviation approach. Since that time the mean-variance approach was extraordinarily successful, even if from time to time other reward-risk combinations may be found in theoretical as well in financial applications literature. A few citations may include Benati-Rizzi, 2007; Huang 2008a, 2008b; and Miller-Ruszczyinsky, 2008.

A third strong analogy is the clear awareness of the difference between efficient (in Paretian sense) solutions and optimal solutions and the idea to concentrate at first on the search for the efficient set of the problem as a preliminary step to any further second stage optimality approach, which in any case should be restricted to the set of efficient solutions. In addition, both authors had quite clear the, now obvious but at that time not so widely known, meaning of efficiency in Paretian sense[4].

[3] On the point see the historical survey provided in Rubinstein, 2006b, p. 102-104

[4] On this topic de Finetti had already written in the 1930s a couple of ground breaking papers, see de Finetti, 1937a, 1937b.

On the basis of these pillars, surprisingly similar despite the certain reciprocal ignorance, both authors framed their problems as a constrained quadratic optimization problem (minimization of the variance of the random gain or rate of return), parameterized to a lower bound constraint on the expected gain E[5], and with the addition of the respective standard constraints.

With regard to the time horizon, both authors distinguished between single period and multiperiod problems. In this paper we neglect the multiperiod question[6] and focus on the analysis of the single period problem, and specifically of the procedure offered by the authors to find the efficient set. Here also the analogies are quite strong and largely prevailing on differences at least as long as we limit our comparison to Markowitz, 1952.

Both authors embedded their feasible sets in the asset space, the n dimensional unitary cube in the reinsurance case and the convex hull of the feasible vertices with respect to the budget constraint in the investment problem. Both analyzed and exploited the geometric character of isomean (parallel planes) and isovariance (concentric ellipsoids) in that space. Both looked at the efficient set as a feasible path in that space, connecting the MfE point of largest (Maximal) feasible Expectation (mean) with the mfV of smallest (minimal) feasible Variance, but to be run in opposite directions (see fig. 1, sect. 10). Both started reasoning on examples in 3-dimensions and both found that the path was a continuous broken line whose (initial, intermediate and final) corner points played a key role. But the ways they arrived to this conclusion were rather different.

In Markowitz the segments of the broken line were part of (half)lines he named "critical lines". Given a subspace of $s \leq n$ assets identified by the

---

[5] In the reinsurance problem this is equivalent to a lower bound constraint on $W+E$ with W a given constant value of the initial guarantee fund of the insurance company.

[6] It was the subject of chapter 11 of Markowitz, 1959 and of the second part of de Finetti, 1940. A comparison between the authors' approach to the multiperiod problem would require another paper. We wish to underline here that Markowitz had a clear idea of the importance of myopia, that is treating any decision as if it were the last. In turn de Finetti found it unconsciously as a byproduct of his strategy to fix retention quotas consistent with an acceptable level of asymptotic ruin probability. The interested reader may find details in Pressacco, 2009, sect. 19.6, p. 527.

budget constraint restricted to the assets of the subspace, and free of individual non-negativity constraints, the critical line of the subspace is the loci of all points (indeed the straight line of tangency between isovariance ellipsoids and isomean planes) with minimum variance for any given level of expectation. He suggested to take the point of minimum feasible variance (mfV) as starting point of the efficient path; this point belongs to a subspace and is obviously a member of the critical line of this subspace. Then, the first segment of the efficient path moves along this critical line in the direction of increasing expectation. Corner points of the efficient path, dictating a change of direction forced by a change of the subspace, may be found for two possible reasons. Either because on the way there is an intersection with a critical line of a larger subspace, which in turn is to be run in the proper direction of increasing mean, or because on the contrary you cannot find any intersection of this type before reaching a stopping point, that is a point where you cannot proceed without breaking one of the individual constraints. Also in this case there is a change of subspace and you move along the critical line of the new smaller subspace (provided that this is at least a two assets subspace, whereas in case of a single asset subspace a "special" rule to choose a new two assets subspace is needed) in the proper direction of increasing mean. The path ends in a final corner where the maximal (largest) feasible expectation (MfE) is reached (which in case of no ties between the assets expectations is a vertex of the convex hull of feasible points; see fig. 2, sect. 10). We suggest to call match (respectively break) corners those of the first (second) type. As we shall see later (sections 6 and 10), there are also "special" corners of a mixed type, coupling at first a break event and then a match event. Variables associated to a match corner (enlarging the space) or to a break corner (lowering the space, or breaking the constraint) will be called match and respectively break variables.

de Finetti took as a starting point the one of largest expectation (MfE), that is full retention of all risks. The efficient path connects through a sequence of segments the MfE point with the endpoint of null variance (zero retention for all risks). The segments are pieces of straight lines (counterpart of critical lines) in a sequence of subspaces (of increasing dimension) corresponding to the subset of risks currently reinsured, while the other are kept fixed at full retention (for a detailed description of the dynamic logic proposed by de Finetti see sect. 4). The risks currently reinsured in a subspace are those sharing the largest advantage from marginal additional reinsurance, measured by the ratio decrease of variance over decrease of expectation and the direction of movement in the subspace is the one on the straight line preserving the equality of the individual advantages (obviously towards decreasing expectation). All intermediate corner points are matching points found where the advantage of beginning to reinsure another risk, previously fully retained, matches from below the decreasing common advantage shared by all risks currently reinsured. Then, according to de Finetti, the efficient path reveals to be a continuous broken line made by *n* segments. In particular, the last segment implies a movement on the *n* dimensional space, with joint additional reinsurance for all risks, and ends in the vertex of null retention. In the Markowitz language, it is clear that each segment of de Finetti's path lies on a critical line of the corresponding subset of assets currently reinsured (but with the other kept at the fixed level $x_i = 1$ of full retention rather than at the level $x_i = 0$ as in the investment problem).

For a better perception of the symmetry between the two approaches, it is convenient to take note of a shortcoming (only recently discovered by Markowitz, 2006) in de Finetti's procedure: he gave for granted that a matching event may be found in any half-line of the efficient path, thus excluding that along some half-line a matching does not happen before a variable reaches its boundary level[7]. In this case, there a break corner would be found, as proceeding in that direction would break an individual constraint. We discuss the point in the final part of sect. 4.

The gap between the two approaches became quite large if we consider the subsequent papers by Markowitz (1956, 1959). Indeed, in the next few years of the 1950s he realized the big jump from the largely informal and intuitive geometric approach previously discussed to the advanced one exploiting the Kuhn-Tucker, 1951, Dantzig *et al*, 1955 and Frank-Wolfe, 1956 results in optimization[8]. This way he was able to generate the formalized sequential procedure known as critical line algorithm (CLA). The same step was never done by de Finetti; only recently and under the decisive stimulus of the Markowitz remark, de Finetti's tools and rules were embedded in the modern environment of mathematical programming to obtain at first a formalized version of the CLA for the reinsurance case (see Pressacco-Serafini, 2007) and later also what could be named a version *a la* de Finetti of the CLA for the standard investment problem (see Pressacco-Serafini, 2009).

[7] Quite likely at 0, but 1 cannot be in our opinion logically excluded as there is no necessary monotonic character in the path of individual retentions, see Pressacco-Serafini, 2007, fig. 1b, p. 39.
[8] Quoted by Markowitz, 2006.

It is interesting to note that Markowitz mimicked de Finetti's choice of the MfE point as starting point of the CLA only in his advanced approach, whereas in the previous intuitive one he had chosen the inverted path direction (starting point mfV).

In addition, it is fair to recognize that de Finetti's version of the CLA requires some additional hypothesis and is then of lesser general applicability than classical Markowitz CLA to cases where non standard constraints hold or there is equality or negativity of assets expectations.

3. de Finetti risk reward approach

de Finetti treated a problem of optimal variable quota share proportional reinsurance faced by an insurance company with an initial guarantee fund (free capital) W and a given portfolio, formally a vector $X = X_1, X_2, \dots, X_n$, of insured risks, with X_i the random liability of the i-th risk. Let $P_i = P(X_i)$ be the net premium (gross premium less the expenses loading for commissions, collection, management, but including the safety loading) charged by the company for the i-th risk and $P = \sum_i P_i$ denote the overall net premium. A proportional retention generated by a variable quota share proportional reinsurance treaty is a feasible vector **x**, i.e. $0 \leq x_i \leq 1$. Under reinsurance on original terms conditions, the insurance company retains a quota x_i and transfers $(1 - x_i)$ of the net premium as well as of the liability to the reinsurer; this way the premium retention is $\Pi = \sum_i \Pi_i = \sum_i x_i P_i$ and the retained liability Y is given by $\sum_i \mathrm{Y}_i = \sum_i x_i X_i$. The post retention random profit of the insurer is then $G = \Pi - Y = \sum_i G_i = \sum_i (\Pi_i - Y_i) = \sum_i x_i (\mathrm{P}_i - X_i)$. How to choose **x**?

At that time, the largely prevailing approach was exclusively driven by the need to keep the default risk (in a single accounting period or over an extended time horizon) conveniently bounded. In a single period setting, the default risk was roughly measured by the probability that $W + G < 0$, or $W < -G$. Unsatisfied by this exclusively risk driven approach, de Finetti suggested an integrated reward-risk approach, based on the simple consideration that every reinsurance reduces the default risk *but leads to renounce to part of the profit* (de Finetti, 1940, p. 2, 2nd paragraph, in the translation Barone, 2006). Then, he looked for a reinsurance strategy able to maximize the *reduction in the risk of default for a given loss in profit* (*ibidem*). This way he explicitly recognized the expected profit (shortly mean) E, and the default or ruin probability RP, as the proper reward and respectively risk measures. After that, de Finetti treated the question as a two-stage problem: in the first stage, the *(E, RP)* Pareto efficient set is selected and, in the second stage, a point within this set is chosen. As strange as it may seem, he did not introduce an explicit formal definition of Pareto efficiency in mean-ruin probability, but there is no doubt that he referred to the standard one: **x** is efficient iff there do not exist other feasible retentions **y** such that both $E(\mathbf{y}) \geq E(\mathbf{x})$ and $RP(\mathbf{y}) \leq RP(\mathbf{x})$, with at least one inequality strict[9].

Under an additional hypothesis of normality of the post retention random gain, another fundamental step in de Finetti's paper was the transition from the single period ruin probability to the single period variance V (or standard deviation σ) of the random gain as the proper risk measure to be used in the first stage of the reinsurance problem. Indeed with $E(\mathbf{x}) = \sum_i x_i m_i = \sum_i x_i (P_i - E(X_i))$ and $V(\mathbf{x}) = \sum_i x_i^2 V(X_i) + 2 \sum_i \quad \sum_{j>i} x_i x_j Cov(X_i, X_j)$, the mean and variance of G, the ruin probability is given by $Prob[(G-E)/\sigma] < [-(W+E)/\sigma]$, or with $t = (W+E)/\sigma$, by $p = Prob\left(\left((G-E)/\sigma\right) < -t\right)$. Note that $t = (W+E)/\sigma$ is surely positive under the hypothesis $m_i > 0$ for any i. Under normality of G (for any **x**) and with N the cumulate of the standard normal (of course a decreasing function of t), it is $p = N(-t)$, a monotone decreasing function of t. Hence, de Finetti argued that minimizing p for any given level of E is equivalent to minimizing σ (see de Finetti, 1940, p. 9, 4th paragraph, *as long as we stay on an iso level W+E constant, determining the maximum of t- i.e. the minimum ruin probability- is equivalent to determining the minimum of σ*).

Although in general, this does not allow us to conclude that the efficient sets *(E,RP)* and *(E,V)* are coincident, this is fortunately true in the reinsurance problem. Hence, he concluded RP and V (or RP and σ) were

[9] It could be unequivocally deduced by his explanation of the formal meaning of his words *maximize the reduction in the risk of default for a given loss in profit* which is given in a footnote (de Finetti, 1940, footnote 2, p. 4, Barone, 2006) as follows: *Precisely, it is an optimum problem, in the sense of my Notes on Problemi di optimum and Problemi di optimum vincolato...*, where the reference is to a couple of enlightening papers (see de Finetti, 1937a, 1937b) concerning the mathematical characterization of the Pareto efficiency set in Economics.

perfectly equivalent risk measures with respect to the problem of finding the single period mean-risk efficient set[10].

After this equivalence result, the problem of finding the efficient reward-risk single period retentions could be formally expressed in the following mean-variance setting:

an insurance company is faced with n risks (policies). The net profit of these risks is represented by a vector of random variables with expected value $\mathbf{m} := \{m_i > 0 : i = 1, \dots, n\}$ and a non-singular covariance matrix $C := \{\sigma_{ij} > 0 : i, j = 1, \dots, n\}$. The company has to choose a proportional reinsurance or retention strategy specified by a retention vector $\mathbf{x}$. The retention strategy is feasible if $0 \leq x_i \leq 1$ for all i. A retention $\mathbf{x}$ induces a random profit with expected value $E = \mathbf{x}^T\mathbf{m}$ and variance $V = \mathbf{x}^T C\mathbf{x}$. A retention $\mathbf{x}$ is by definition mean-variance efficient or Pareto optimal if, for no feasible retention $\mathbf{y}$ we have both $\mathbf{x}^T\mathbf{m} \leq \mathbf{y}^T\mathbf{m}$ and $\mathbf{x}^T C\mathbf{x} \geq \mathbf{y}^T C\mathbf{y}$, with at least one inequality strict.

4. de Finetti's procedure to find the mean-variance efficient set

After this theoretical background, de Finetti looked for a procedure to find the single period *(E,V)* efficient retentions set. Working at a time where constrained programming and Kuhn-Tucker conditions were things to come, he applied an approach that is quite close to a dynamic programming framework, conveniently supported by geometric intuitions. Precisely, he suggested to look for such a set X^* as a path in the n dimensional unitary cube of feasible retentions. The path starts at the full retention vertex **x**=**1**, point of largest expectation (being $m_i > 0$ for any i) and ends at the opposite, full reinsurance or zero retention, vertex **x**=**0**, unique point of minimum variance, thanks to the non singularity of *C.* At any point x^* of the

[10] The event of ruin in a single period could be seen as equivalent to the disaster event which has been the base of the celebrated Roy, 1952 approach to portfolio selection. Inspired by a safety first strategy, Roy looked for the risky portfolio which minimizes the probability of a single period disaster and showed that it is the portfolio that minimizes the ratio $(E - d)/\sigma$, where $(E - d)$ is the excess return with respect to the disaster return d. This strategy is currently widely applied in a lot of decision problems including financial ones. See for example Huang, 2008a and Dorfleitner-Utz, 2012.

path an efficient movement is made in such a way to get locally the largest benefit measured by the ratio decrease of variance over decrease of expectation. To translate this idea in an operational setting, de Finetti introduced tools and rules. The tools are a set of advantage functions:

$$F_i(\mathbf{x}) = \frac{1}{2}\frac{\frac{\partial V}{\partial x_i}}{\frac{\partial E}{\partial x_i}} = \sum_{j=1}^{n} \frac{\sigma_{ij}}{m_i} \cdot x_j, \qquad i=1,\dots,n \tag{1}$$

designed to capture (half of) the local benefit coming at **x** from an additional or initial reinsurance of the *i*-th risk.

The rules give an answer to the following three questions:

1. to which risks should we provide additional reinsurance at $\mathbf{x}^*$?
2. if this set is not a singleton, which feasible direction should we choose? and
3. where and why should we change direction?

The answer to the first question is: at any point $\mathbf{x}^*$ of X^* move in such a way to provide (additional or initial) reinsurance only to the set $I(\mathbf{x}^*)$ of those risks sharing the largest value of $F_i(\mathbf{x}^*)$ (put $max_i F_i(\mathbf{x}^*) = \lambda(\mathbf{x}^*)$). If this set is a singleton, the direction of the efficient path is obvious; otherwise, the answer to the second question is: move in the direction preserving the equality of all the $F_i(\mathbf{x}^*)$ coming from $I(\mathbf{x}^*)$. Given the form of the advantage functions, it is easily seen that this implies a movement on a segment of the cube characterized by the set of equations $F_i(\mathbf{x}) = \lambda(\mathbf{x})$, for *i* belonging to $I(\mathbf{x}^*)$. As the direction of movement is driven exclusively by the risks belonging to $I(\mathbf{x}^*)$ we call them directional risks at $\mathbf{x}^*$. The chosen direction is kept up until a change in the *I* set happens. This gives, according to de Finetti, the following answer to the third question: change direction at the first point on the segment where a risk *i*, previously not belonging to the *I* set, matches from below the common value of the advantage functions of the currently directional risks, thus beginning to be reinsured and becoming a new member of the *I* set. But this implies the addition of one equation to the system giving the best direction of movement and, as a consequence, a change of direction that corresponds to a corner point in the efficient path. de Finetti was then able to conclude that the efficient *(E,V)* path is made by a continuous broken line of *n* segments, whose corners (except for the terminal vertex) correspond to points in which a new risk begins to be reinsured.

Moreover, there is a one to one monotonic correspondence between the efficient set X^* and the interval of values of the advantage parameter $0 \le \lambda \le F_1(\mathbf{1}) = max_i F_i(\mathbf{1})$.

More formally, a sequence of corner values λ_h, corner points $\mathbf{x}_h$ and sets I_{h-1} of directional risks in the interval $\lambda_h < \lambda \le \lambda_{h-1}$ is described by the equations $\lambda = F_i(x)$ for $i \in I_{h-1}$, $x_i = 1$ for $i \notin I_{h-1}$, which define $\mathbf{x}$ parameterized by λ. The corner point $\mathbf{x}_h$ is still found in this direction, where one of the non-directional risks (previously frozen at the level 1 of full retention) matches from below the common value of the advantage functions of the *h-1* directional risks in I_{h-1}. There, the new matching risk joins the set I_{h-1} of the previous directional risks, changing state and defining the new set I_h.

This sequential procedure is to be seen as the informal predecessor of the famous critical line approach applied by Markowitz to find *(E,V)* efficient portfolios of financial assets. de Finetti gave for granted that matching points are in a one-to-one correspondence with segments of the efficient path.

In addition, de Finetti found closed-form formulae for the efficient retentions $\mathbf{x}(\lambda)$ in case of no correlation (see de Finetti, 1940, p. 12). Indeed, he gave an outline of the way to obtain the same result in a particular case of correlation structure (group correlation, see de Finetti, 1940, pp. 28-29). He also underlined that these nice results came from the possibility to define an *a priori* fixed ordering of (entrance in reinsurance) the risks, while, in the general case, the application of the sequential procedure previously described is needed.

A hidden bug in de Finetti's approach went unnoticed until it was recently discovered by Markowitz (2006) in his review of de Finetti's paper. The point is that de Finetti's rules have only an internal coherency but lack an unconditional one. Indeed, it was implicit in de Finetti's reasoning, and a necessary and sufficient condition for the survival of the whole procedure, that a matching event could be found in any segment of the efficient path. Let us shortly denote it by matching hypothesis (MH). But we cannot be sure that MH holds for a general structure of correlation, except for the special cases of no correlation or group correlation. Then, Markowitz, while recognizing de Finetti's primacy in the (*E*,*V*) approach, expressed the opinion that his (de Finetti's) procedure works only in the special case of no correlation. *For the case of uncorrelated risks, de Finetti solved the problem of computing the set of mean variance efficient portfolios. He explains while the problem with correlated risks is more complicated and solves special cases* (the group correlation) *of it* (see Markowitz, 2006, p.5). *He does not solve it in general for correlated risks* (see Markowitz, 2006, p.11), while for the general case of correlation it is necessary to keep account of the constraint binding at the break event and to apply advanced mathematical programming techniques.

From a strict technical point of view this is undisputable, yet we think that a softer sentence would offer a fair historical resume of de Finetti's contribution. Indeed, on one side, de Finetti's procedure may quite likely work in a lot of real world correlation structures of insurance portfolios. On the other side, also in case MH does not hold, de Finetti's logic, properly adjusted to purge it from its incoherency, remains an enlightening way to manage and solve the efficient retention problem. Last but not least, if we approach that problem with the modern tools of constrained quadratic programming, the optimality conditions can be restated in the most natural form through the advantage functions so that the revised version of de Finetti's approach is fully compatible with the modern technology (on the point see Pressacco-Serafini, 2007, pp. 34-36).

5. Cases of no correlation or group correlation: some results

In the case of no correlation (see de Finetti, 1940, p. 12) $F_i(\mathbf{x})$ reduces to $F_i(\mathbf{x}) = (x_i V_i / m_i) = x_i v_i$. There is a natural ordering of risks induced by $v_1 > v_2 > \cdots > v_n$ and the efficient set may be expressed in closed form formulae. For any value $0 = \lambda(\mathbf{0}) \leq \lambda \leq \lambda(\mathbf{1}) = v_1$, an efficient retention is given by $x_i(\lambda) = min(\lambda / v_i\,; 1)$, *i=1,…,n*; conversely $\lambda(x) = max_i x_i v_i$ so that a one-to-one monotone correspondence between λ and **x** is established.

Putting λ values on the abscissa, the graph of the functions $x_i(\lambda)$ are broken lines made by two segments, the first one going out from the origin with slope $1/v_i$ which at $\lambda = v_i$ becomes the horizontal line at level 1.

After this analysis de Finetti was able (see de Finetti, 1940, p. 12) to express both *E* and *V* of any efficient retention as a function of λ. Precisely, in the interval $\lambda_h < \lambda \leq \lambda_{h-1}$ (*h-1* risks already reinsured), it is:

$$E(\lambda) = \sum_{i=1}^{h-1} x_i m_i + \sum_{i=h}^{n} m_i = \lambda \sum_{i=1}^{h-1} (m_i / V_i) m_i + \sum_{i=h}^{n} m_i \qquad (2)$$

and putting $A_{h-1} = \sum_{i=1}^{h-1} m_i^2 / V_i$ e $B_{h-1} = \sum_{i=h}^{n} m_i$, we get:

$$E(\lambda) = \lambda A_{h-1} + B_{h-1} \qquad (3)$$

$$V(\lambda) = \sum_{i=1}^{h-1} x_i^2 V_i + \sum_{i=h}^{n} V_i = \lambda^2 \sum_{i=1}^{h-1} (m_i^2 / V_i^2) V_i + \sum_{i=h}^{n} V_i =$$

$$= \lambda^2 \sum_{i=1}^{h-1} m_i^2 / V_i \; + \sum_{i=h}^{n} V_i \qquad (4)$$

and with $C_{h-1} = \sum_{i=h}^{n} V_i$, it is:

$$V(\lambda) = \lambda^2 A_{h-1} + C_{h-1} \qquad (5)$$

It turns out that, in the no correlation case, *E* and *V* are piecewise linear and respectively piecewise quadratic functions of λ. Moreover, it is easy to show that they are continuous but not differentiable at the connection points (corner points of the efficient path). Unfortunately de Finetti neglected to analyze the properties of the efficient set in a *(E,V)* reference system; in particular he did not exploit the above relations to obtain *V* as the following immediate piecewise quadratic function of *E*:

$$V = C_{h-1} + ((E - B_{h-1})^2 / A_{h-1}) \qquad (6)$$

By doing so he could have had at his disposal an easy road to show that, also at the connection points, *V* is a continuous and differentiable function of *E* with, not surprisingly, $\partial V / \partial E = 2\lambda$. Note that λ could be interpreted as a shadow price of insurance and $F_i(\mathbf{x})$ as marginal utility at $\mathbf{x}$ of further reinsurance of the i-th risk, so that, given λ, reinsurance is provided up to the point where decreasing marginal utility equals the shadow price[11].

By group correlation following de Finetti, who at that time considered this structure a good proxy of the real insurance world, we mean that the risks are partitioned into a number g of groups, $q=1,\ldots,g$ each one characterized by a couple (l_q,ρ_q) of constants. Denoting by m_{iq} and σ_{iq} the expectation and the standard deviation of the i-th risk of the group q, l_q is a group specific loading coefficient used to charge net insurance premiums through a safety loading principle inspired by the standard deviation principle so as $m_{iq}=l_q\sigma_{iq}$[12]. The constant $\rho_q>0$ is a group specific correlation coefficient and plays a role in specifying the covariance matrix C, which is a block diagonal matrix $(C_1,\ldots,C_g)$, with non null elements only on the main diagonal squared blocks given, for $i_q \neq j_q$, by $\rho_q\sigma_{iq}\sigma_{jq}$.

Under this structure, the advantage functions became, with x_q the retention vector of the group q:

$$F_{iq}(\mathbf{x}_q) = l_q^{-1}\left(x_{iq}\sigma_{iq} + \sum_{j_q\neq i_q} x_{jq}\sigma_{jq}\right) \quad (7)$$

de Finetti argued that, under this hypothesis, some of the nice properties of the no correlation case still hold and this would allow a quasi closed-form solution to the problem. He gave only a sketch of how to get this result. It comes from the possibility to define an *a priori* ordering of risks within each group (based on the standard deviation ranking $\sigma_{1q}>\sigma_{2q}>\ldots>\sigma_{nq}$), as well as an *a priori* ordering of groups (based on the advantage function ranking of the first risk of each group at full retention $F_{11}(\mathbf{1})> F_{12}(\mathbf{1})>\ldots> F_{1g}(\mathbf{1})$).

Indeed, applying these suggestions and following de Finetti's procedure, we were able elsewhere (Pressacco-Ziani, 2012, p. 352) to check that the matching condition really holds, and that the value of the advantage

---

[11] de Finetti's paper had a large impact on the actuarial sciences, even if the applications remained restricted to the no correlation case. On the point see Bühlmann- Gerber, 1978; Gerber, 1984; Gerber-Shiu, 2003.

[12] In this section, the first index refers to a risk and the second to a group and the symbol must not be confused with notations like σ_{ij} which means covariance between the two risks i and j.

parameter at which the i-th risk of the group q starts to be reinsured, is given by:

$$\lambda_{iq} = l_q^{-1}\left[\sigma_{iq}\left(1+\rho_q(i-2)\right)\right] + \sum_{j=2}^{n_q}\sigma_{jq} \qquad (8)$$

and that closed-form formulae hold for the efficient retentions both in the retention space (see Pressacco-Ziani, 2012, p. 352) and in the mean-variance one (for details, Pressacco-Serafini-Ziani, 2011, pp. 444-448).

In particular, we found that in the *(E,V)* plane the efficient set is still piecewise parabolic continuous and differentiable, also at the connection points. That is enough to show that de Finetti's logic could surely work also under (admittedly very special) conditions of non null correlation. Besides that, we suggest that this model could be considered a counterpart of the simplified models of covariance that (in order to make the problem computationally more manageable) have been developed in the wave of portfolio selection and CAPM models (Markowitz, 1999 quoted Sharpe, 1963, Cohen-Pogue, 1977 and Elton-Gruber, 1973).

6. Adjusting de Finetti's procedure for the case of break points

de Finetti's procedure may be adjusted also when the MH does not hold (for details see Pressacco-Serafini, 2007, discussion in sect. 5 as well as the example in sect. 8). Here we just summarize the results.

It turns out that the corner points may be of three different types: match (in turn distinct between below match and above match), break and mixed corners. At a below match corner the advantage function of a risk previously fully retained matches from below the current value of the advantage parameter, joins the set of directional risks and begins to be reinsured; at an above match corner the advantage function of a risk previously fully reinsured matches from above the current value of the advantage parameter, joins the set of directional risks and ceases to be fully reinsured; at a break corner the retention level of a previously partially reinsured risk reaches a boundary value (quite likely 0, but 1 cannot be excluded), leaves the set of directional risks and is (may be temporarily) frozen at the boundary level; at a mixed corner (surely corresponding to a vertex of the unitary cube) there is at first a break when the retention level of the unique variable previously partially reinsured reaches the lower boundary 0, and leaves the set of directional risks which as a consequence becomes empty. Then, in order to leave the vertex, we must wait for a subsequent matching between the decreasing value of the advantage parameter and the highest value at the vertex of the advantage functions of the currently fully retained risks. Except in case the efficient set contains one or more vertices (mixed corners) the one to one correspondence between the interval $0 \leq \lambda \leq F_1(\mathbf{1})$ and the vector of efficient retentions $\mathbf{x}^*(\lambda)$ is preserved.

In the *(E,V)* plane the efficient set is given by a continuous union of parabolic arcs corresponding to the segments of the broken line. The graph is continuous also at the connection points and differentiable everywhere (except if the efficient set includes one or more vertices as there are kinks at the connection points of the graph corresponding to such vertices). A convenient hypothesis of non degeneracy is implicitly assumed in the above description. It may be resumed by the condition that at any transitional value of the advantage parameter λ, only one risk is in transition.

We signal that Markowitz pointed out clearly (Markowitz, 2006) that, if de Finetti would have completed a careful analysis of the problem without relying too much on the wrong intuition of the absence of break corners, he could have made a giant step toward the Kuhn-Tucker conditions (Kuhn-Tucker, 1951).

7. Advantage functions in the investment problem

In the next sections, we will see how to extend de Finetti's approach to a standard portfolio selection problem. Let us recall the essential of such a problem.

An investor is faced with *n* assets. The net rate of return of these assets is represented by a vector of random variables with expected value $\mathbf{m} := \{m_i : i = 1, \dots, n\}$ and a non-singular covariance matrix $C := \{\sigma_{ij} : i, j = 1, \dots, n\}$. In addition to the usual non degeneracy (singularity) conditions on the covariance matrix, we will assume a labeling of the assets coherent with a strict ordering of expectations, namely $m_1 > m_2 > \cdots > m_n$. The investor has a budget, conveniently normalized to 1, to invest in the given assets. Let x_i be the fraction of budget invested in the asset i. If short positions are not allowed, the portfolio strategy is feasible if $x_i \geq 0$ for all i and $\sum_i x_i = 1$. A portfolio $\mathbf{x}$ induces a random rate of return with expected value $E = \mathbf{x}^T \mathbf{m}$

and variance $V = \mathbf{x}^T C\mathbf{x}$. A feasible portfolio **x** is by definition mean-variance efficient or Pareto optimal if for no feasible portfolio **y** we have both $\mathbf{x}^T\mathbf{m} \leq \mathbf{y}^T\mathbf{m}$ and $\mathbf{x}^T C\mathbf{x} \geq \mathbf{y}^T C\mathbf{y}$, with at least one inequality strict. Let X^* be the set of optimal portfolios.

de Finetti did not treat at all (neither in his 1940 paper nor even later) the asset portfolio problem; yet we found that de Finetti's procedure, once the behind logic is transferred from the reinsurance to the investment problem, remains a good strategy. To obtain in a natural and straightforward way something analogous to the critical line algorithm, we make recourse to a modified version of the advantage functions tailored to the constraints of the investment problem.

Now, the role of the "elementary" advantage function is played by F_{ij}, i.e. the advantage associated to the portfolio adjustments coming from a "small" trading between asset *i* (decreasing) and asset *j* (increasing).

More formally and with reference to a portfolio **x** with positive quotas of both assets *i* and *j*, let us call "bilateral" trading in the *i-j* direction the portfolio adjustment obtained through a "small" exchange between *i* decreasing and *j* increasing. If the benefit (burden) measure is given by the ratio decrease (increase) of variance over decrease (increase) of expectation, the advantage functions ought to be defined as[13]:

[13] Note that the sign of the denominator is the same as *(j-i)*; and also that, as both *i* and *j* are active assets (that is with positive quotas) in the current portfolio **x**, a feasible bilateral trading may happen also in the *j* - *i* direction. Then both $F_{ij}(\mathbf{x})$ and $F_{ji}(\mathbf{x})$ describe the results of a feasible bilateral trading at **x**. Moreover it is immediate to check that $F_{ij}(\mathbf{x}) = F_{ji}(\mathbf{x})$. Yet, it is convenient to think that the economic meaning of the two functions is symmetric: precisely, if without loss of generality *i* is less than *j*, $F_{ij}(\mathbf{x})$ describes a benefit in algebraic sense, while $F_{ji}(\mathbf{x})$ describes a burden. If in the current portfolio x_i is positive and x_j is null, then the only feasible bilateral trading may be in the direction *i-j*. And $F_{ij}(\mathbf{x})$ describes a benefit if *i* is less than *j* or a burden in the opposite case. Obviously, if both x_i and x_j are at level 0 no feasible trade between *i* and *j* may take place.

$$F_{ij}(\mathbf{x}) := \frac{1}{2}\frac{\frac{\partial V}{\partial x_i}-\frac{\partial V}{\partial x_j}}{\frac{\partial E}{\partial x_i}-\frac{\partial E}{\partial x_j}} = \sum_{h=1}^{n} \frac{\sigma_{ih}-\sigma_{jh}}{m_i-m_j} x_h \tag{9}$$

8. A procedure for the standard investment problem under regularity

Let us now use the advantage functions to build a friendly procedure to find the optimum mean-variance set for the standard portfolio problem.

The starting point of the mean-variance path is $\mathbf{x}_2^* = (1,0,0,\dots,0)$, which is the point with largest expectation due to the ordering convention. The choice of the index 2 may sound strange, but it is justified as it denotes that a second asset starts to be active at this point. Indeed, we leave $\mathbf{x}_2^*$ in the direction granting the largest benefit, that is the largest value over $j=2,\dots,n$ of

$$F_{1j}(\mathbf{x}_2^*) := \frac{\sigma_{11}-\sigma_{j1}}{m_1-m_j} \coloneqq \lambda(\mathbf{x}_2^*) \tag{10}$$

Let us label asset j_2 this asset. This means that the bilateral trading in the direction 1- j_2 gives the largest benefit $\lambda(\mathbf{x}_2^*)$ and dictates the efficient path leaving $\mathbf{x}_2^*$ in the direction $(-\varepsilon, 0,0,0,\varepsilon,\dots 0)$. While the trade takes effect, the $\mathbf{x}$ values change and consequently $F_{1j}(\mathbf{x}) \coloneqq \lambda(\mathbf{x})$ decreases. In the regular case, the bilateral trading 1- j_2 remains the most efficient one until we find a point on the above segment. There the benefit granted by this trade is matched (and regularity means just that such a feasible matching point may be found before the end of the segment) by another bilateral trade, that is until the nearest point (labelled $\mathbf{x}_3^*$), where $F_{1j_2}(\mathbf{x}_3^*) = F_{1j}(\mathbf{x}_3^*) = \lambda(\mathbf{x}_3^*)$ for some $j \neq j_2$. Let us label asset j_3 this asset. Some remarks are in order now.

Remark a): at $\mathbf{x}_3^*$ also the bilateral trade j_2-j_3 matches the same benefit $\lambda(\mathbf{x}_3^*)$. Indeed the following result holds: for any triplet i, j, h of assets and any portfolio $\mathbf{x}$ such that at least x_i and x_j are strictly positive,

$$F_{ij}(\mathbf{x}) = F_{ih}(\mathbf{x}) = \lambda \Rightarrow F_{ih}(\mathbf{x}) = \lambda \tag{11}$$

If $i < j < h$, the advantage functions value describes a matching of the benefits given by any bilateral trade *i-j, i-h, j-h*. The proof is straightforward.

Remark b): small feasible changes in a portfolio composition may come as well from joint movements of more than two assets, to be seen as a multilateral trade. However, any multilateral feasible trade may be defined as a proper combination of feasible bilateral trades and if all the bilateral trades share the same benefit, then the benefit of the multilateral trade too matches the common benefits. This explains why we may concentrate on the benefits from bilateral trades neglecting the analysis of the consequences of multilateral ones.

Remark c): in some cases, it could be advantageous to split a multilateral trade in bilateral components, implying also one or more trading of the type *j-i* with $j>i$. Surely in isolation this cannot be a drift of movements along the efficient path. However, as we shall see later, it could add efficiency when inserted in the context of a multilateral trade.

Let us go back now to the point $\mathbf{x}_3^*$ where all bilateral trades between two of the first three assets and hence any feasible multilateral trade among them shares the same efficiency. At first sight, there is here an embarrassing lot of opportunities to leave $\mathbf{x}_3^*$ along directions granting the same benefit. However, help is given by the second rule we receive in heritage from the reinsurance problem (see Section 4): move along the path (indeed the segment) which preserves the equality of all implied advantage functions. In our case move along the direction $(\varepsilon_1, 0{,}0, \varepsilon_{j_2}, \ldots, \varepsilon_{j_3}, 0, \ldots 0)$ (with $\varepsilon_1 + \varepsilon_{j_2} + \varepsilon_{j_3} = 0$) which preserves the $F_{1j_2}(\mathbf{x}) = F_{1j_3}(\mathbf{x})$ equality until a new matching point is found.

Under regularity, that is if a feasible matching point may be found at each step, it is easy to prove (see Appendix) that a sequential application of this procedure defines a piecewise linear optimal path with corner points $\mathbf{x}_h^*$ where the matching asset henceforth labeled j_h joins the other assets already active in the portfolio.

The ending point of the efficient path is the point $\mathbf{x}_p^*$ of absolute minimum variance with benefit $\lambda(\mathbf{x}_p^*) = \mathbf{0}$ for all advantage functions.

To summarize, in case of regularity, corner points of the optimum path are always matching points and each matching point $\mathbf{x}_h^*$ corresponds to a matching of the advantage function $F_{1j_h}(\mathbf{x}^*)$ of a newcomer asset j_h with the common value $\lambda(\mathbf{x}^*)$ of the advantage functions $F_{1i}(\mathbf{x}^*)$ of the current active assets i. Yet only in the didactic but fully unrealistic case of no correlation, the labeling induced by the expectation vector is coincident with the entrance (matching) ordering (except of course for the asset of largest expectation, which is still associated with the starting point). In addition, except for the no correlation case, we cannot a priory but only ex post say if there is regularity or not. We remark that under regularity the computational burden is quite lower than what may appear at first glance. Indeed, there is no need to compute values of the advantage functions for all pairs *i-j* (at least *i* active), but it is enough to evaluate the *n-1* values $F_{1j}(\mathbf{x})$. Here we may say that the first asset (asset 1) plays overall the optimum path the role of reference variable. More generally, it is required that the current referent variable is active. And the optimality condition to be checked is that there is a non-negative λ such that, for all active assets i, $F_{1i}(\mathbf{x}) = \lambda$, while, for the other non-active j, $F_{1j}(\mathbf{x}) \leq \lambda$, with strict equality holding only at corner points just for the newcomer matching asset. As for the other feasible values of $F_{ij}(\mathbf{x})$ and their superfluous role in checking the optimality conditions the following result holds:

- both *i* and *j* active then $F_{ij}(\mathbf{x}) - \lambda = 0$
- only i active then $F_{1i}(\mathbf{x}) = \lambda$ and $F_{1j}(\mathbf{x}) < \lambda \Rightarrow \big(F_{ij}(\mathbf{x}) - \lambda\big)(j - i) < 0$

This explains that the optimality conditions at **x** require simply that all basic bilateral tradings between each pair of active assets share the same benefit level λ, while all basic tradings between an active *i* and a non-active *j* have a lower efficiency level[14].

[14] More precisely less efficiency means lower benefit *(i<j)* or greater burden *(j>i)*. At matching corners the matching variable becomes efficient and even if it is at the moment still non-active shares the efficiency of all previous active assets.

9. The standard investment problem under non-regularity

Let us now pass to treat the non-regular case. The non-regularity comes from the fact that there is a failure in the matching sequence in the sense that along a segment of the optimum path one of the active assets reaches its boundary value 0 before a matching occurs. This is the counterpart of a break event in the reinsurance case and it should be clear that at a break point the break variable leaves the efficient set and remains frozen, maybe temporarily, at the level 0.

The new set of the other active assets (given that there are at least 2), determines the new direction, either preserving the equality between the advantage functions in case of three or more active assets, or in the unique feasible direction if there are only two assets. Before discussing what happens in case of only one surviving active asset, we underline that the behaviour at a break corner is the only difference between the non-regular and the regular case. As to the computational burden, it is still enough to compute at any point of the optimum path only the values of *(n-1)* advantage functions.

Yet an additional effort may be required at those break points, where the break asset has been playing the role of the reference asset. Indeed its exit from the set of active assets requires the choice of a new reference (the reference must be active) and then the need to compute the values of a new set of *(n-1)* advantage functions. Hereafter the usual conditions of efficiency still hold for the new set of advantage functions.

Let us finally discuss what happens at a break point where the set of active assets is a singleton, so that the point is a vertex of the n dimensional simplex of feasible allocations. This being the case, a resetting of the procedure is in order. Denoting by k the original label of the singleton asset, we look for leaving the vertex in such a way as to maximize the efficiency of a bilateral trade of the *k-j* type. This means looking for the largest positive value of $F_{kj}(\mathbf{x})$ over all $j>k$. The corresponding value of the benefit parameter could be seen as a match value. We may say that such vertex corners are mixed corners, that is points where both a break (at first for a larger λ) and a match event (later for a smaller λ) happen. At these corners the one-to-one correspondence between efficient portfolio and values of λ is lost: there is an interval of λ values to which the same corner points $\mathbf{x}$ is associated. Consequently, the graph of the function *V(E)* in the mean-variance space has a kink. For details see the example in sect. 10.

10. Comparison between de Finetti and Markowitz

Let us comment on the affinity and differences between Markowitz and de Finetti's oriented procedure as regards the use of geometric intuitions in the solution of the investment problem. They inspired the paper of Pressacco-Serafini, 2009 in which an extension of de Finetti's logic to the investment problem is offered (resumed here in sections from 7 to 9) and were a pillar of the Markowitz earliest paper (Markowitz, 1952) and of chapter 7 (entitled Geometric analysis of efficient sets, pp. 129-153) of his subsequent book (Markowitz, 1959). In the first paper, he analyzed in detail the three asset case in the standard version with non negativity constraints $x_i \geq 0$, and one collective budget constraint $\sum_i x_i$. Exploiting the budget constraint the feasible set is represented in a bidimensional setting (in the plane (x_1, x_2)) by the rectangular triangle whose vertices are the points (0,0), (0,1), (1,0). Then, for a while, he went to what he subsequently (Markowitz, 1987, p. 39) called Black model, whose only constraint is the budget one; then exploiting it in the form $x_3 = 1 - (x_1 + x_2)$, transformed the Black model in a free problem in 2-dimensions. He showed that, in the plane (x_1, x_2), isomean lines are parallel straight lines whose equation (adding here the additional hypothesis of no ties between assets expectations) is:

$$x_2 = \frac{(m-m_3)}{(m_2-m_3)} - x_1 \frac{(m_1-m_3)}{(m_2-m_3)} \tag{12}$$

while isovariance curves are a set of concentric ellipses of equation:

$$V = x_1^2(V_1 + V_3 - 2\sigma_{13}) + x_2^2(V_2 + V_3 - 2\sigma_{23}) + 2x_1x_2(\sigma_{12} + V_3 - \sigma_{13} - \sigma_{23}) + 2x_1(\sigma_{13} - V_3) + 2x_2(\sigma_{23} - \sigma_{23}) + V_3 \tag{13}$$

i.e.

$$V = ax_1^2 + bx_2^2 + 2cx_1x_2 + 2dx_1 + 2ex_2 + f \tag{14}$$

whose center is the point of (unconditional) minimum absolute Variance (maV) with coordinates: $x_1 = (ce - bd)/(ab - c^2)$, $x_2 = (cd - ae)/(ab - c^2)$. Then, he argued that the set of portfolios of minimum variance for any fixed level of mean (geometrically points of tangency between the corresponding isomean and isovariance lines) is a straight line whose equation is given in implicit form by:

$$(m_2 - m_3)(ax_1 + cx_2 + d) = (m_1 - m_3)(bx_2 + cx_1 + e) \qquad (15)$$

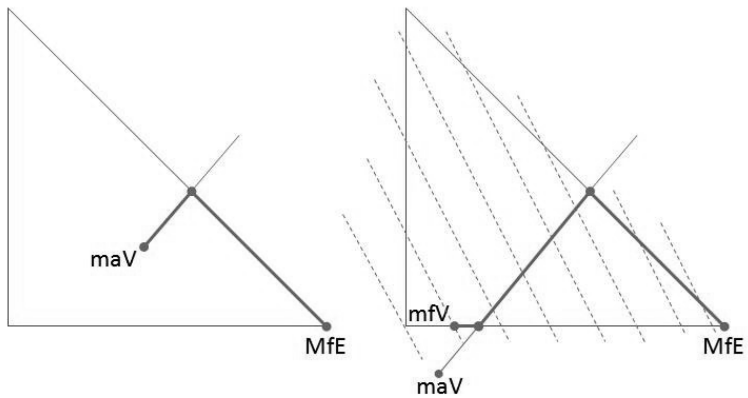
He defined critical line such a line, which includes of course the maV point. The half-line starting from maV and going in the direction of increasing mean is the set of efficient mean-variance points (portfolios) of the Black problem in the 2-dimensional picture. This could be considered the main critical half line in the three asset space, but other critical (half)lines should be considered in the two asset (sub)spaces (on the direction of the sides of the triangle) and even in the one asset (sub)spaces (single points). In the intuition of Markowitz, the critical lines played a key role to find the efficient mean variance set. Indeed he suggested (see Markowitz, 1952, p. 87, footnote 10) the following dynamic vision of the efficient path: *the efficient set may be traced out by starting at the point of minimum feasible variance moving continuously along various subspaces according to definite rules...typically we proceed along a given critical line until either this line intersects one critical line of a larger subspace or meets a boundary and simultaneously the critical line of a lower dimensional subspace. In either of these cases the efficient line turns and continues along the new line. The efficient line terminates when a point with maximum feasible mean is reached.*

In other words, the fundamental result is that the efficient path is a continuous broken line made by a sequence of segments lying on critical lines of (sub)spaces so as corner points of the efficient path correspond to an intersection of two adjacent critical (half)lines.

To see more in detail how this procedure should work, it is convenient to discuss a couple of examples. In the first one (analyzed in Markowitz, 1952, p. 85) the maV is internal to the triangle (hence feasible), while the MfE is the vertex (1,0,0) of the triangle (see fig. 1 left side).

Figure 1: Efficient path (bold face segments): first (left) and second (right) case. Horizontal axis: x_1. Vertical axis: x_2. Diagonal side: $x_1 + x_2 = 1$ i.e. $x_3 = 0$. Assets are labelled according to their expectations $(m_1 > m_2 > m_3)$, so that isomean are parallel straight lines with negative slope greater, in absolute value, of the one of the diagonal side of the feasible triangle.

The first segment of the efficient path (left side) lies on the critical half-line of the three asset space; the path has a (break) corner at the point where this critical line intersects a side (the diagonal) of the feasible triangle; then the efficient path goes on the critical half-line of the corresponding two asset subspace ending in the MfE vertex.



According to de Finetti's point of view, the path would start at the MfE vertex of largest expectation, would move in the direction of the largest trading advantage, i.e. the diagonal side of the triangle, where trading between asset 1 and asset 2 is the most efficient (so that portfolios made by the first two assets are efficient), up to the corner identified by a matching of the advantage obtained by trading also with the third asset. Here a new direction of movement along the direction preserving the equality of the advantage of all bilateral tradings, implies portfolios with positive quotas of all three assets and the procedure ends in the point where there is no longer any advantage in further bilateral trading, which is the maV point.

It is clear that what is a subspace with *s* assets in Markowitz, is equivalent to trading involving *s* assets in de Finetti's logic; according to the first point of view, the critical line is the loci of smallest variance for any level of feasible expectation in the subspace, in the other approach it is the direction of largest advantage by trading with those assets. The inversion of the path direction implies also an inversion of the interpretation of match and break corners in the two procedures. In this example, there is a unique intermediate corner; in Markowitz it is found where the critical line of the 3 asset space intersects the boundary (diagonal side) of the triangle. In this direction this ought to be interpreted as a break event, because going on in that direction would imply to break one of the constraints: $x_3 \geq 0$. Conversely, in de Finetti's logic at that corner there is a match between the advantage of trading between assets number 1 and 2 and the one coming from involving in trade also asset number 3: a match event.

In the other example[15] treated by Markowitz (1952, p. 86), the maV is not feasible, but part of the critical half line of the 3 asset space is feasible (internal to the triangle). Here the mfV is on the basis side of the triangle, so the first segment is on the critical half line of the corresponding 2 assets subspace. A first intermediate corner is found where this line intersects the critical line of the 3 assets space (to be seen as a match event) and a second and last intermediate corner comes when this critical line intersects the diagonal side of the triangle (a break event). Finally, the last part of the efficient path is the segment lying on the critical line of the corresponding two assets subspace to reach once more the MfE vertex.

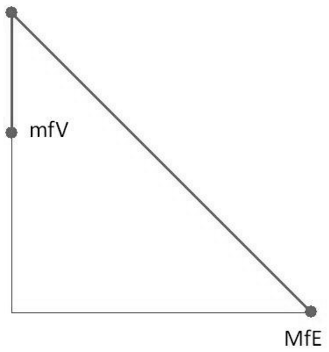
In de Finetti's logic (starting point at MfE, right side of fig. 1) the first segment corresponds to the last one of Markowitz, the first intermediate corner is driven by a match (like the one of the previous example), the second intermediate corner is found when a break event happens (proceeding in that direction would be breaking the short selling constraint for the asset number 2). Hereafter, the asset number 2 is no more part of efficient portfolios and the last segment directs, through the most efficient constrained trading between asset 1 and asset 3, toward the end point of mfV where further trading is no more advantageous. Summarizing, the ordering of segments and corners in de Finetti's and Markowitz geometric driven procedures are exactly opposite as well as the match and break events.

In another example (Markowitz, 1959, p. 143), Markowitz discussed a case in which the main critical half line is fully external to the feasible triangle, so that the efficient path is exclusively made by segments on two sides of the triangle including the vertex (0,1,0), which he defined intermediate corner with interesting features (see fig. 2). *The set of efficient portfolios seems to reverse direction at the vertex. Nevertheless, it always*

[15] In order to have the same ordering of expectations in both cases, our labelling of assets is different from the one adopted by Markowitz in the second example.

moves in the direction of increasing return. Another interesting feature will be noted later in connection with the relationship between expected return and variance of return of efficient portfolios. This remark concerns the behaviour of the graph of efficient portfolios in mean-variance space and not in the asset one. Markowitz underlined (p. 153) *that to each segment on a critical line of the asset space there corresponds a parabola in the mean variance space and that at the intersection of two successive critical lines the typical relation between the corresponding parabolas is that they are tangent: they touch but do not cross* (said another way the graph of $V(E)$ is continuous and differentiable also at connection points), there are no kinks. *It is however possible for the curve to have a kink when the set of efficient portfolios turns a vertex corner.*

Figure 2: Efficient path: an intermediate vertex corner.



This shows that Markowitz had already in the 1950s a quite clear idea on kinks of the efficient set in the mean-variance space. On the point he offered sometimes later (see Markowitz, 1987), chapter 11 entitled "Kinks in the set of efficient EV combinations") a clear and conclusive proof involving the advantage function as revealed by the following sentence: *there is a kink only if there is a discontinuity of the function dV/dE at the intersection point of two consecutive critical lines*. And it is clear from the following discussion (see Markowitz, 1987, p. 259) that a kink occurs in the standard case only at a vertex of the feasible polytope or if all the assets currently in portfolio have the same mean.

At that time, this was not an original result; a few years before a paper by Dybvig (1984) had given the end to the debate on kinks on the *(E,V)* space that, notwithstanding the clear ideas of the father of the *EV* approach, remained still open 25 years later the cited Markowitz remark (Markowitz, 1959). In the abstract of the paper (Dybvig, 1984, p. 239), after having defined switching points in the mean-variance frontier those corresponding to changes in the set of assets held, Dybvig recalls that *traditional wisdom holds that each switching point corresponds to a kink, while Ross (1977) has claimed that kinks never occur! The paper shows that the truth lies between the two views, since the efficient frontier may or may not be kinked at a switching point. There is some indication that kinks are rare, since a kink corresponds to a portfolio in which all active assets have the same expected return. In particular as an immediate corollary if all securities have different expected returns a kink may occur only if a single security is held there*, (Dybvig, 1984, p. 243). Proofs were based on arguments regarding properties of the mean-variance frontier.

We wish to underline here that de Finetti's approach offers, at least under the assumption of no ties in assets expectations, the most simple way to understand the reasons why kinks may occur only in an intermediate vertex corner of the asset space: only at these mixed corners there is a sudden downward jump in the value of the advantage parameter λ corresponding to (half) the ratio derivative of variance over derivative of expectation on the efficient path, hence only at these points there is a jump in the first derivative of the function $V(E)$ (see the last paragraph in sect. 9).

This easy deduction is nothing but a byproduct of the new insights about the fundamental properties of the efficient set, both in the asset and in the mean variance space, allowed by (the adjusted version of) de Finetti's approach also in the investment problem.

To conclude it is fair to recall that de Finetti never treated in detail the behaviour of efficient portfolios in a EV space; on the contrary Markowitz was able to immediately recognize the importance of this reference space, which became of overwhelming importance in the subsequent development of portfolio theory and capital markets equilibria.

11. Conclusions

It is today plainly recognized that de Finetti introduced the mean variance approach in financial problems under risk and, dealing with a problem of proportional reinsurance, offered a correct procedure to find the mean variance efficient set in case of no correlation and, under a convenient regularity assumption (no break points), also in case of correlation. In this paper we concentrate on a couple of additional insights provided by de Finetti's approach. As regards the reinsurance problem, we explain how a natural adjustment of his procedure, fully respecting his logic and applying his tools (the advantage functions) and rules, provides correct solutions also in case of non regularity (in case of breaks too). Then, with reference to the standard portfolio selection problem, we show that, exploiting a convenient modification of the advantage functions tool, the extension of this adjusted procedure to the standard portfolio selection problem is able to offer an alternative way to look at the set of efficient portfolios and a new clear characterization of its properties in terms of benefit from asset trading. In particular, as shown in the example discussed in sect. 10, this approach offers a simple, yet insightful, explanation of when and why corner points appear in the path of efficient portfolios in the asset space, and clearly describes the properties of different types of corner points. Moreover, it offers an equally simple explanation of the properties at the corner points of the graph of efficient portfolios in the mean-variance space.

Appendix: a mathematical programming formulation for the standard investment problem.

In the appendix (based on sect. 19.7 Pressacco-Serafini, 2009) we provide a mathematical foundation of the approach through advantage functions illustrated in sections 7-9. The problem we investigate can be restated as the following quadratic problem:

$$\min \frac{1}{2} \mathbf{x}^T C \, \mathbf{x}$$

$$\mathbf{m}^T \mathbf{x} \geq \mathrm{E}$$

$$\mathbf{1}^T \mathbf{x} = \mathbf{1}$$

$$\mathbf{x} \geq 0 \tag{16}$$

for every attainable E, i.e. $\min_i m_i \leq E \leq \max_i m_i$. The strict convexity of the objective function guarantees that there is a one-to-one correspondence between points in X^* and optimal solutions of (16) for all attainable E such that the constraint $\mathbf{m}^T\mathbf{x} \geq E$ is active. The Karush-Kuhn-Tucker conditions are necessary and sufficient for optimality of (16), since the constraints are regular and the objective function is strictly convex (see Shapiro, 1979; Karush, 1939; Kuhn and Tucker, 1951). The conditions are expressed through the Lagrangean function:

$$L(\mathbf{x}, \lambda, \mu, \mathbf{v}) = {}^1\!/_2\, \mathbf{x}^T C\mathbf{x} + \lambda(E - \mathbf{m}^T\mathbf{x}) + \mu(1 - \mathbf{1}^T\mathbf{x}) - \mathbf{v}\mathbf{x} \tag{17}$$

and state that $\mathbf{x}^*$ is optimal if and only if there exist Lagrange multipliers $(\lambda^*, \mu^*, \mathbf{v}^*)$, $\lambda^* \geq 0$, $\mathbf{v}^* \geq 0$, such that:

1) $\mathbf{x}^*$ minimizes $L(\mathbf{x}, \lambda^*, \mu^*, \mathbf{v}^*)$
2) $\mathbf{x}^*$ is feasible in formulae (16)
3) either $x_j^* = 0$ or $v_j^* = 0$ (or both) and
   either $\mathbf{m}^T\mathbf{x} = \mathrm{E}$ or $\lambda^* = 0$ (or both) (18)

In order to verify 1) of (18), since $\mathbf{x}$ is unconstrained (in the Lagrangean minimization), it is enough to compute:

$$\partial L/_{\partial x} = C\mathbf{x} - \lambda\mathbf{m} - \mu\mathbf{1} - \mathbf{v} = \mathbf{0} \tag{19}$$

i.e., componentwise

$$\textstyle\sum_j \sigma_{hj}x_j - \lambda m_h - \mu - v_h = 0, \qquad h = 1, \dots, n \tag{20}$$

We want to rephrase the optimality condition by showing how the optimal variables depend on λ. They depend also on μ, but we prefer to hide this dependence by solving (19) first on μ. We assume that the indices are ordered as $m_1 > m_2 > \mathrm{K} > m_n$. Let k be any index such that $x_k > 0$. We denote this variable as the reference variable. We have $v_k = 0$ by complementarity and, for *h=k,* formula (20) for *h=k* is:

$$\textstyle\sum_j \sigma_{kj}x_j - \lambda m_k - \mu = 0 \tag{21}$$

Now we subtract (20) from (21):

$$\sum_j (\sigma_{kj} - \sigma_{hj}) x_j - \lambda (m_k - m_h) + \nu_h = 0, \qquad h \neq k \qquad (22)$$

or equivalently:

$$\frac{\sum_j (\sigma_{kj} - \sigma_{hj}) x_j}{m_k - m_h} + \frac{\nu_h}{m_k - m_h} = \lambda \qquad h \neq k \qquad (23)$$

Observe that solving (23) is equivalent to solving (20). Indeed, once (23) is solved, μ can be easily computed from all other variables. We have defined the advantage functions as:

$$F_{kh}(\mathbf{x}) = \frac{\sum_j (\sigma_{kj} - \sigma_{hj}) x_j}{m_k - m_h} \qquad h \neq k \qquad (24)$$

(note that $F_{kh}(\mathbf{x}) = F_{hk}(\mathbf{x})$) and by using the advantage functions we may rephrase (23) as:

$$F_{kh}(\mathbf{x}) + \frac{\nu_h}{m_k - m_h} = \lambda \qquad h \neq k \qquad (25)$$

Now we partition the variable indices $\{1, \ldots, n\} \setminus k$ into three sets as:

$$I_k^* = \{h \neq k : x_h > 0\}, I_0^* = \{h < k : x_h = 0\}, I_k^0 = \{h > k : x_h = 0\} \qquad (26)$$

For the sake of notational simplicity, we omit to denote that these subsets actually depend also on $\mathbf{x}$. Moreover, let $I^* := I_k^* \cup \{k\}$ and $I^0 := I_0^k \cup I_k^0$ (respectively the sets of positive and null variables independently of the reference variable). Then, taking into account that $\nu_h \geq 0$, $m_k > m_h$, if $h \in I_k^0$ and $m_k < m_h$, if $h \in I_0^k$, the complementarity condition can be restated through the advantage functions in the following form:

Optimality condition. Let k such that $x_k > 0$. Then $\mathbf{x} \geq 0$ is optimal if and only if $\mathbf{1}^T \mathbf{x} = 1$ and there exists $\lambda \geq 0$ such that:

$$F_{kh}(\mathbf{x}) \geq \lambda, h \in I_0^k, F_{kh}(\mathbf{x}) = \lambda, h \in I_k^*, F_{kh}(\mathbf{x}) \leq \lambda, h \in I_k^0 \qquad (27)$$

The following facts can be easily deduced from the optimality condition:

Corollary. Let $i \in I_k^*$. If $j \in I_k^*$ then $F_{ij}(x) = \lambda$. If $j \in I^0$ then

$$F_{ij}(\mathbf{x}) \leq \lambda \text{ if } i > j, \;\; F_{ij}(\mathbf{x}) \geq \lambda \text{ if } i > j \qquad (28)$$

This result implies that the role of reference variable can be subsumed by any variable in I_k^* without changing the optimality conditions, provided the sets I_k^*, I_0^k and I_k^0 are duly redefined according to the new reference, and, more importantly, for the same value of λ. In other words, resetting the reference does not affect the value of λ.

The set I_k^* can be empty only in the extreme cases $x_k = 1$ and $x_h = 0$, $h \neq k$. In this case (27) becomes:

$$F_{kh}(\mathbf{x}) = \frac{\sigma_{kk}-\sigma_{hk}}{m_k-m_h} \leq \lambda, h > k \text{ and } F_{kh}(\mathbf{x}) = \frac{\sigma_{kk}-\sigma_{hk}}{m_k-m_h} \geq \lambda, h < k \quad (29)$$

Hence, if:

$$\max\left\{\max_{h>k} \frac{\sigma_{kk}-\sigma_{hk}}{m_k-m_h} ; 0\right\} \leq \min_{h<k} \frac{\sigma_{kk}-\sigma_{hk}}{m_k-m_h} \quad (30)$$

the point $x_k = 1$ and $x_h = 0$, $h \neq k$, is optimal with λ taking any value within the above interval. Note that the point $(1,0,\ldots,0)$ is always optimal (since the r.h.s. term is missing) and that $x_k = 1$ and $x_h = 0$, $h \neq k$, can be optimal only if $\sigma_{kk} < \sigma_{hk}$ for all $h<k$ (necessary but not sufficient condition). In particular, the point $(0,0,\ldots,1)$ of absolute minimum mean can be also mean-variance efficient if and only if $\sigma_{nn} < \sigma_{hn}$ for all h. In this case it is the end point of the set X^*.

If the set I_k^* is not empty, the optimality condition $F_{kh}(x) = \lambda$, $h \in I_k^*$, is a linear system in the variables in I^*:

$$\sum_{j \in I^*} \frac{\sigma_{kj}-\sigma_{hj}}{m_k-m_h} x_j = \lambda, \qquad h \in I_k^* \quad (31)$$

Adding the condition $\sum_{j \in I^*} x_j = 1$ yields a square linear system whose solution is an affine function of λ:

$$x_h := w_h + \lambda z_h, \qquad h \in I^* \quad (32)$$

(with w solution of the linear system with r.h.s. $(0,0,\ldots,0,1)$ and z solution with r.h.s. $(1,1,\ldots,1,0)$) and clearly $x_h = 0$, $h \in I^0$.

As stated in Pressacco-Serafini (2007) the minimum portfolio variance in (16) is a strictly convex monotonically increasing function of the mean E and the multiplier λ is its derivative (or a subgradient on the points of non-differentiability). Therefore the set X^* can be parameterized via λ instead of E, taking into account that some points of X^*, where the derivative of the function has a discontinuity jump, correspond to an interval of values for λ. It is easy to understand that this discontinuity jump corresponds to the so-called kinks of the parabolic efficient frontier in the mean-variance space.

Basing on the advantage functions we could obtain a computational procedure, analogous to the critical line algorithm by Markowitz, to describe X^* parameterized via λ.

References

Barone, L. (2006), The problem of "full risk insurances", *Journal of Investment Management*, 4, 19-43, English translation of Chapter 1 of de Finetti 1940

Benati, S. and R. Rizzi (2007), A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem, *European Journal of Operational Research*, 176-1, 423-434

Bühlmann, H. and H. Gerber (1978), Risk bearing and the reinsurance market, *The ASTIN Bullettin*, 10, 12-24

Cohen, K.J. and J.A. Pogue (1977), An empirical evalutation of alternative portfolio selection models, *Journal of Business*, 40-2, 166-193

Dantzig, G., A. Orden and P. Wolfe (1955), The generalized simplex method for minimizing a linear form under linear inequality restraints, *Pacific Journal of Mathematics*, 5-2, 183-195

de Finetti, B. (1937a), Problemi di optimum, *Giornale Istituto Italiano Attuari*, 8, 48-67

de Finetti, B. (1937b), Problemi di optimum vincolato, *Giornale Istituto Italiano Attuari*, 8, 112-126

de Finetti, B. (1940), Il problema dei pieni, *Giornale Istituto Italiano Attuari*, 9, 1-88

de Finetti, B. (1969), Un matematico e l'economia, *Franco Angeli*, Milano

Dorfleitner, G. and S. Utz, (2012), Safety first portfolio choice based on financial and sustainability returns, *European Journal of Operational Research*, 221-1, 155-164

Dybvig, P. (1984), Short sales restrictions and kinks on the mean variance frontier, *The Journal of Finance*, 39, 239-244

Elton, E.J. and M.J. Gruber (1973), Estimating the dependence structure of share prices, *Journal of Finance*, 28-5, 1203-1232

Frank, M. and P. Wolfe (1956), An algorithm for quadratic programming, *Naval Research Logistic Quarterly*, 3, 95-110

Gerber, H.U. (1984), Equilibria in a proportional reinsurance market, *Insurance, Mathematics and Economics,* 3, 97-100

Gerber, H. U. and E.S W. Shiu (2003), Economic Ideas of Bruno de Finetti in the Wiener Process Model, *Metodi Statistici per la Finanza e le Assicurazioni*, B.V. Frosini Editor,75-95

Huang, X. (2008a), Portfolio selection with a new definition of risk, *European Journal of Operational Research*, 186, 351-357

Huang, X. (2008b), Mean-semivariance models for fuzzy portfolio selection, *Journal of computational and applied mathematics* 217-1, 1-8

Karush, W. (1939), *Minima of functions of several variables with inequalities as side constraints*, M.Sc. dissertation, Department of Mathematics, University of Chicago, Chicago, IL, USA

Kuhn, H.W. and A.W. Tucker (1951), Nonlinear programming, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman ed., University of California Press, Berkeley, CA, USA

Markowitz, H. (1952), Portfolio selection, *Journal of Finance*, 6, 77-91

Markowitz, H. (1956), The optimization of quadratic functions subject to linear constraints, *Naval Research Logistics Quarterly*, 3, 111-133

Markowitz, H. (1959), Portfolio selection: efficient diversification of investments. New York: John Wiley & Sons. (reprinted by Yale University Press, 1970; 2nd ed. Basil Blackwell, 1991)

Markowitz, H. (1987), Mean Variance analysis in portfolio choice and capital markets, Basil Blackwell Inc. Cambridge.

Markowitz, H. (1999), The early history of portfolio theory, *Financial Analyst Journal*, 55-4, 5-16

Markowitz, H. (2006), de Finetti scoops Markowitz, *Journal of Investment Management*, 4, 5-18

Miller, N. and A. Ruszczynski (2008), Risk-adjusted probability measures in portfolio optimization with coherent measures of risk, *European Journal of Operational Research*, 191-1, 193-206

Pressacco, F. (1986), Separation theorems in proportional reinsurance, *Insurance and Risk Theory*, M. Goovaerts et al. eds., Reidel Publishing, 209-215

Pressacco, F. (2009), Bruno de Finetti, actuarial science and the theory of finance in the 20th century. *Vinzenz Bronzin's option pricing models*, Springer Verlag, 519-534

Pressacco, F. and P. Serafini (2007), The origins of the mean-variance approach in finance: revisiting de Finetti 65 years later, *Decisions in Economics and Finance*, 10-1, 19-49

Pressacco, F. and P. Serafini (2009), New insights on the mean-variance portfolio selection from de Finetti suggestions, New Frontiers in Insurance and Bank Risk Management, McGraw Hill Milano, 253-270

Pressacco, F., P. Serafini and L. Ziani (2011). Mean-Variance efficient strategies in proportional reinsurance under group correlation in a Gaussian framework. *European Actuarial Journal*, 1-2, 433-454

Pressacco, F. and L. Ziani (2012). The influence of correlation and loading on MV efficient retentions in variabile quota share proportional reinsurance. *Mathematical and Statistical Methods for Actuarial Sciences and Finance*. Eds. Cira Perna e Marilena Sibillo, MAF Series, Springer Verlag, 349-358.

Ross, S. (1977), The capital asset pricing model, short-sale restrictions and related issues, *Journal of Business*, 32, 177-183

Roy, A.D. (1952), Safety first and the holding of assets, *Econometrica*, 20, 431-449

Rubinstein, M. (2006a), Bruno de Finetti and mean-variance portfolio selection, *Journal of Investment Management*, 4, 3-4

Rubinstein M. (2006b), *A history of the theory of investments*, Wiley, New York

Shapiro, J.F. (1979), *Mathematical programming: structures and algorithms*, Wiley, New York

Sharpe, W. (1963), A simplified model for portfolio analysis, *Management Science*, 9-2, 277-293

Frank Lad* - Giuseppe Sanfilippo†

SCORING ALTERNATIVE FORECAST DISTRIBUTIONS: COMPLETING THE KULLBACK DISTANCE COMPLEX

Abstract

We develop two surprising new results regarding the use of proper scoring rules for evaluating the predictive quality of two alternative sequential forecast distributions. Both of the proponents prefer to be awarded a score derived from the other's distribution rather than a score awarded on the basis of their own. A Pareto optimal exchange of their scoring outcomes provides the basis for a comparison of forecast quality that is preferred by both forecasters, and also evades a feature of arbitrariness inherent in using the forecasters' own achieved scores. The well-known Kullback divergence, used as a measure of information, is evaluated via the entropies in the two forecast distributions and the two cross-entropies between them. We show that Kullback's symmetric measure needs to be appended by three component measures if it is to characterise completely the information content of the two asserted probability forecasts. Two of these do not involve entropies at all. The resulting "Kullback complex" supported by the 4-dimensional measure is isomorphic to an equivalent vector measure generated by the forecasters' expectations of their scores, each for one's own score and for the other's score. We foreshadow the results of a sophisticated application of the Pareto relative scoring procedure for actual sequential observations, and we propose a standard format for evaluation.

JEL CLASSIFICATION: C10, C11

KEYWORDS: Total Logarithmic Scoring Rule; Prevision; Entropy/Extropy; Cross Entropy; Pareto Optimal Exchange; Kullback Symmetric Divergence; Bregman Divergence.

* University of Canterbury, Department of Mathematics and Statistics, Christchurch, New Zealand, *E-mail address:* F.Lad@math.canterbury.ac.nz

† Department of Mathematics and Computer Science, University of Palermo, Italy, *E-mail address:* giuseppe.sanfilippo@unipa.it

1. Introduction

Bruno de Finetti's "partly-baked" musings with I. J. Good (1962) on the use of a penalty function to gauge the relative predictive qualities of competing forecast distributions have developed by now into a complete theory of proper scoring rules, replete with practical computational procedures. Gneiting and Raftery (2007) present a fairly comprehensive technical review of both the theory and its application. Lad (1996, Chapter 6) presents a pedagogical introduction. Proper scoring procedures are meant to provide the basis for a subjectivist understanding of forecast evaluations that can replace completely the common practice of "hypothesis testing." Such testing is regarded by subjectivists to be without foundation. The very notion that data are generated by an unobservable probability distribution is rejected in their characterisation of probability and its connection with empirical statistical research. In this present tribute to the imagination of de Finetti we propose a striking development in scoring theory. Our analysis gives rise to a completed understanding of Kullback's information measure in the context of the total logl scoring function. Implications for statistical practice are substantial.

The logarithmic scoring function for a probability mass vector (pmv) $\mathbf{p}_N$ asserted by a forecaster "p" for a quantity X on the basis of an observation $X = x^o$ within its realm of possibility $\mathscr{R}(X) = \{x_1, x_2, ..., x_N\}$ is specified as the logarithm of the probability assessed for the outcome x^o that is actually observed: $S(X = x^o, \mathbf{p}_N) \equiv log(p^o)$. It has been known since early developments of proper scoring rules by Savage (1971) and by Matheson and Winkler (1976) that a pmv proponent's expectation of one's own log score equals the negentropy value embedded in that pmv. Symbolically, $E_p[log\, P(X = x^o))] = \sum p_i\, log\, p_i$. Moreover, p's expectation of the amount by which p's own log score will exceed the score of someone else's assertions, $\mathbf{q}_N$, is equal to the relative entropy in $\mathbf{p}_N$ with respect to $\mathbf{q}_N$. That is,

$$E_p[S(X, \mathbf{p}_N) - S(X, \mathbf{q}_N)] = \sum p_i\, log\, (\frac{p_i}{q_i}).$$

This relative entropy measurement is also widely known as the Kullback-Leibler (1951) directed divergence between the pmv's $\mathbf{p}_N$ and $\mathbf{q}_N$, and is commonly denoted by $D(\mathbf{p}_N || \mathbf{q}_N)$. A non-symmetric function of the two pmv's, the value of $D(\mathbf{p}_N||\mathbf{q}_N)$ is summed with its counterpart $D(\mathbf{q}_N||\mathbf{p}_N)$ to yield Kullback's (1959) symmetric distance measure, denoted by $\mathscr{D}(p||q) \equiv D(\mathbf{p}_N||\mathbf{q}_N) + D(\mathbf{q}_N||\mathbf{p}_N)$. The analysis and the applications of this distance function in information theory and practice have been extensive. Among recent theoretical developments, the Kullback-Leibler function is now known to constitute a "Bregman divergence" relative to the Bregman function $\Phi(\mathbf{p}_N) = \sum p_i \, log \, p_i$, studied by Censor and Zenios (1997). This structure sets it in a context shared by several other divergence functions, though not by still others. The recent contribution of Dawid and Musio (2015) references how lively and widespread is research interest in related issues.

Subsequent consideration of scoring rules by Lad, Sanfilippo and Agrò (2012) has addressed the fact that the log score for mass functions, while proper, is incomplete in that it does not make use of *all* the observations with their asserted probabilities that are entailed in the assertion of $\mathbf{p}_N$ and the observation of X. Augmented by the sum of $(1-p_i)\,log(1-p_i)$ values for possibilities of X that *are* observed *not* to obtain, the log scoring rule becomes the *total log scoring rule*. Its prevision (expectation) according to the proponent who asserts $\mathbf{p}_N$ equals the negentropy *plus* the negextropy of the asserted mass function. Negextropy was characterized in an article by Lad, Sanfilippo, and Agrò (2015) as a complementary dual of negentropy. The structure of the Kullback-Leibler divergence entertains a complementary dual as well: $D^c(\mathbf{p}_N||\mathbf{q}_N) = \sum(1-p_i) log(\frac{1-p_i}{1-q_i})$. Parallel considerations define the *complementary* symmetric Kullback distance as $\mathscr{D}^c(\mathbf{p}_N||\mathbf{q}_N) \equiv D^c(\mathbf{p}_N||\mathbf{q}_N) + D^c(\mathbf{p}_N||\mathbf{q}_N)$. Together, the Kullback distance and its complementary distance sum to the Kullback total symmetric distance: $\mathscr{D}_T(\mathbf{p}_N||\mathbf{q}_N) \equiv \mathscr{D}(\mathbf{p}_N||\mathbf{q}_N) + \mathscr{D}^c(\mathbf{p}_N||\mathbf{q}_N)$.

The present article presents two new developments in our understanding of associated issues. In the first place, we consider the expectations of the asserters of both $\mathbf{p}_N$ and $\mathbf{q}_N$, each for one's own total log score *and* for the other's score. Attention turns to these proponents' considerations of the net gains they each might achieve on the basis of their professed willingness to trade their achieved scores for their personally assessed prevision values. (This constitutes the very definition of their previsions.) We find an appropriate strategy for assessing the relative merits of the two pmv assertions: perhaps surprisingly, it is to award to each forecast proponent the *net gain achieved by the other's* pmv, scaled appropriately, rather than to award to each forecaster one's own achieved proper score.

Secondly, our analysis turns to the pair of expected net gains and expected comparative gains of the forecasters. We learn that Kullback's total symmetric distance measure can be generated equivalently by two other pairs of functions that are different from that which specifies the Kullback-Leibler divergence. This result suggests that Kullback's symmetric distance function should be extended to a four-dimensional vector function if it is to characterise all the distinct features of the information expressed in the two pmv's. It turns out that the vector space spanned by this 4-D vector function is isomorphic to the space spanned by the two proponents' previsions for their own scores and for each other's score.

Section 2 of this article identifies the two pmv proponents' expected *net gains* and expected *comparative gains* according to their uncertain assessments of an unknown quantity X. Section 3 develops the logic of the Pareto-optimal assessment of their comparative forecast quality we propose here, while the *scale* of this comparative scoring procedure is discussed in Section 4. Scaling requires consideration of the variances assessed by the two forecasters for various score components. Attention turns to the higher dimensions of the Kullback complex that arise from these considerations in Section 5, and Section 6 displays why the appended dimensions of the complex are *not* Bregman divergences. Section 7 portrays the historical context of Kullback's analysis of the problem, and outlines its reconsideration according to the subjectivist understanding. Section 8 presents briefly some experience with and some proposals for application to statistical problems.

Deliberations in this article make use of the syntax preferred for Bruno de Finetti's operational subjective construction of coherent prevision, discussed by de Finetti (1967, 1972) and Lad (1996, pp. 39-43). Events are characterised as numerical quantities (numbers, not sets) with the realm of possibility $\mathscr{R}(E) = \{0,1\}$. Otherwise they are addressed with the same syntax as any other quantity whose realm is finite and discrete. Prevision, $P(X)$, is a personal value assertion about X, a price at which its proponent is avowedly indifferent to a prize of X or of $P(X)$. Since a prevision is an assertion made by someone, the P symbol shall be subscripted by a mark denoting who is making the assertion, as in $P_p(\cdot)$ and $P_q(\cdot)$ used here. We shall refer to the proponents of these previsions merely as "p" and "q", respectively. One's prevision for an event is also called a probability. Parentheses around an expression that may be true or false denotes an event whose numerical value equals 1 if the expression is true, and 0 if it is false. Thus, for example, $E_i \equiv (X = x_i)$ is an event that equals 1 if and only if it happens that X is in fact found to equal x_i; otherwise $E_i = 0$. In the more commonly used measure-theoretic syntax of formalist probability ala Kolmogorov, prevision is equivalent to expectation; in these terms, the expectation of a $0-1$ random variable is a probability.

2. Interpersonal assessment of the Total Log Scores

The purpose of this Section is to formalise the notation for the total logarithmic score that will be used to assess our two forecasting distributions (pmvs), and to introduce the formalities of two concepts that will be central to our considerations. The first is the *net* gain achieved by a forecaster who asserts a pmv for an observable quantity. It defined as the difference between the score that will be achieved and the forecaster's expectation for the score. The second is the *comparative* gain achieved by one of the forecaster's score relative to that achieved by the other. It is defined simply as the difference between their two scores. We can then identify each forecaster's professed expectation for each of these gains. In doing so we will have identified *four fundamental previsions* that are relevant to evaluating the forecast quality of the two forecast pmv's on the basis of observations.

Consider the standard problem of statistical forecasting an unknown quantity X whose realm of possible observation values is $\mathscr{R}(X) = \{x_1, x_2, ..., x_N\}$. An associated vector of partition events, $\mathbf{E}_N$, is composed of its logically related components $E_i \equiv (X = x_i)$ for $i = 1, 2, ..., N$. The sum of these events must equal 1 since one and only one of them must equal 1. Because the considered value of X is unknown when the forecasters assess probabilities, it is just not known which of the events is the 1 and which of them are the 0's. This will be identified when we observe the value of X. We shall study the uncertain opinions of two people who publicly assert distinct probability mass functions via the vectors $\mathbf{p}_N$ and $\mathbf{q}_N$, respectively. Once the value of X is observed, their assertions will be assessed via a widely touted proper score, the Total Log Scoring function, defined for $\mathbf{p}_N$ (and similarly for $\mathbf{q}_N$) by

$$\begin{aligned} S_{TL}(X, \mathbf{p}_N) &\equiv \sum_{i=1}^{N} E_i \, log(p_i) + \sum_{i=1}^{N} \tilde{E}_i \, log(1-p_i) \\ &= log(p^o) + \sum_{i:\, p_i \neq p^o} log(1-p_i) \\ &= log(\frac{p^o}{1-p^o}) \; + \; \sum_{i=1}^{N} log(1-p_i) \; , \end{aligned} \tag{1}$$

as long as the pmv $\mathbf{p}_N$ is strictly positive and $N > 1$. Here p^o denotes the probability asserted for the value of $(X = x^o)$ that eventually comes to be observed; and the sum in the second line of (1) includes all summands of the form $log(1-p_i)$ *except for* $log(1-p^o)$.

The "total log score" is a negative valued score, but it constitutes an award. A larger score (smaller absolute value) is achieved by a more informative distribution. As extremes, the pmv that accords probability 1 to the value of X that occurs, $(X = x^o)$, receives a score of 0; those which accord probability 1 to any particular value of X that does *not* occur receive $-\infty$. The uniform pmv receives $log(1/N) + (N-1)\, log(1-1/N) \approx -log(N) + e^{-1}$ for large N.

N.B. From this point on we shall no longer write subscripts on vector quantities to denote their size, nor the limits on the range of summations, always understood to be summations over units from 1 to N.

To foreshadow a relevant consideration that will be addressed at the very end of the analysis, we note here that each of the two summands constituting the total log score, $\sum p \; log \, p$ and $\sum (1-p)\; log(1-p)$ is itself a proper score of the pmv $\mathbf{p}$, as is any positive linear combination of these two. Whereas the constructions we now discuss are developed in terms of the total log score, they pertain to each of these addends separately in the same way.

2.1. Previsions for one's own score and for another's score

In the following discussion we shall denote the entropy and extropy functions, respectively, by $H(\mathbf{p}) = -\sum p\, log(p)$ and $J(\mathbf{p}) = -\sum(1-p)\, log(1-p)$. Sums such as $-\sum p\ log(q)$ and $-\sum q\ log(p)$ are called "cross entropies", and will be denoted by $CH(\mathbf{p}||\mathbf{q})$ and $CH(\mathbf{q}||\mathbf{p})$. Similarly, "cross extropies" will be denoted by $CJ(\mathbf{p}||\mathbf{q}) \equiv -\sum(1-p)\ log(1-q)$. Recognise that cross entropies/extropies are not symmetric in their arguments. For example, $CH(\mathbf{p}||\mathbf{q}) \neq CH(\mathbf{q}||\mathbf{p})$.

Theorem 1. The linearity of coherent prevision identifies the previsions of both p and q for their own scores and for the other's score as:

$$\begin{aligned}
P_p[S_{TL}(X,\mathbf{p})] &= \sum p\, log(p) + \sum(1-p)\, log(1-p) = -[H(\mathbf{p}) + J(\mathbf{p})]\ , \text{ and} \\
P_p[S_{TL}(X,\mathbf{q})] &= \sum p\, log(q) + \sum(1-p)\, log(1-q) = -[CH(\mathbf{p}||\mathbf{q}) + CJ(\mathbf{p}||\mathbf{q})]\ , \\
P_q[S_{TL}(X,\mathbf{q})] &= \sum q\, log(q) + \sum(1-q)\, log(1-q) = -[H(\mathbf{q}) + J(\mathbf{q})]\ , \text{ and} \\
P_q[S_{TL}(X,\mathbf{p})] &= \sum q\, log(p) + \sum(1-q)\, log(1-p) = -[CH(\mathbf{q}||\mathbf{p}) + CJ(\mathbf{q}||\mathbf{p})]\ .
\end{aligned}$$

The proof is immediate from the application of linear coherent previsions to the first line of equation (1).

These four previsions exhaust the content of what p and q assert about their own prospective score values and for the scores to be achieved by one another, before the value of X is observed. We shall designate them as *the four fundamental previsions*, and we shall learn why this nomenclature is appropriate. Let's think a moment about their relative sizes.

Since the scoring function $S_{TL}(\cdot,\cdot)$ is proper, each of p and q expects to achieve a greater score than will the other; for the proponent of a pmv expects to achieve a maximum score by publicly professing one's actual pmv, **p**, rather than any other pmv, **q**. For this reason, the use of proper scoring rules is said to promote honesty and accuracy in the profession of probabilities. Thus, for example, $-[H(\mathbf{p}) + J(\mathbf{p})]$ surely exceeds $-[CH(\mathbf{p}||\mathbf{q}) + CJ(\mathbf{p}||\mathbf{q})]$, a statement which is true when **p** and **q** exchange places as well. However, the relative sizes of the two assessors' own expected scores, $P_p[S_{TL}(X,\mathbf{p})]$ and $P_q[S_{TL}(X,\mathbf{q})]$, may be in any order. This just depends on what they each assert about X, which is **p** or **q**, respectively.

When cross-score previsions are considered, p's prevision for q's score may be *either smaller or larger* than q's expectation for q's own score. In a computational survey of pmv's within the unit-simplex we found that the size of $P_p[S(X,q)]$ exceeds $P_q[S(X,q)]$ in about 10 to 15 percent of paired $(\mathbf{p}_3, \mathbf{q}_3)$ selected uniformly at random from the unit simplex. We have not yet characterised the situations in which this occurs, though it should be straightforward.

N.B. From this point we no longer use bold print to distinguish vectors from their components. These will be identifiable by their context.

We are now ready to analyse important features of p's and q's expectations regarding one another's scoring performance.

2.2. Expected Net Gains and Expected Comparative Scores

Consider first the "net gains" to be achieved by the proponents of p and q as a result of their receiving the total log scores, $S(X,p)$ or $S(X,q)$ respectively, as a gain in return for paying out the prices at which each of them values this score:

$$\begin{aligned} NG(X,p) &\equiv S(X,p) - P_p[S(X,p)] \ , \ \text{and} \\ NG(X,q) &\equiv S(X,q) - P_q[S(X,q)] \ . \end{aligned} \tag{2}$$

Both p and q expect a personal net gain of 0, since their assertions are presumed to be coherent. One's prevised (expected) score is the price at which a person values this unknown score to be achieved. The proponent of a pmv is avowedly willing both to buy and to sell a claim to the score for this price. For example, p is avowedly indifferent to the values of $S(X,p)$ and $P_p[S(X,p)]$. Thus, $P_p[NG(X,p)] = 0$, as does $P_q[NG(X,q)] = 0$.

However, the proponents of p and q do *not* expect each others' net gains to equal 0. Before assessing them, it is worthwhile to recognize that there is something arbitrary about our definition of Net Gain. For the forecaster is also indifferent to an award of the negative net gain, the result of an exchange in which the forecaster offers to sell the value of $S(X,p)$ in return for receiving

$P_p[S(X,p)]$. This would reverse the sign of an awarded Net Gain as we have defined it. This is the sense in which an award to p of a proper score value or its prevision assessment would be arbitrary. Nonetheless, we are now ready to study it as defined.

Theorem 2. $P_p[NG(X,q)]$ may be positive or negative valued, as may $P_q[NG(X,p)]$.

As proof, note specifically that

$$\begin{aligned} P_p[NG(X,q)] &= P_p[S(X,q)] - P_q[S(X,q)] = \\ &= -[CH(p||q) + CJ(p||q)] + H(q) + J(q) \\ &= \sum (p-q)\, log(\frac{q}{1-q})\ ; \\ \text{while} & \\ P_q[NG(X,p)] &= P_q[S(X,p)] - P_p[S(X,p)] = \\ &= -[CH(q||p) + CJ(q||p)] + H(p) + J(p) \\ &= \sum (q-p)\, log(\frac{p}{1-p})\quad . \end{aligned} \tag{3}$$

Now we have already noticed and remarked that p's prevision for q's score may be greater than or less than q's prevision for the same. If p expects a score for q greater than q expects, then p will expect a positive net gain for q because p is thinking that q expects too low a score. In such a case,

$$P_p[NG(X,q)] = P_p[S(X,q)] - P_q[S(X,q)] \ > \ P_q[S(X,q)] - P_q[S(X,q)] = 0.$$

If p expects a score for q less than q expects, the reverse is be true: $P_p[NG(X,q)] = P_p[S(X,q)] - P_q[S(X,q)] \ < \ 0$.

The forecasters' expectations of their *comparative* gains relative to the other's are another matter. Defining the "comparative gain of p over q" on the basis of observing X as

$$CG(X,p,q) \quad \equiv \quad S(X,p) - S(X,q) \ = \ -CG(X,q,p)\ , \tag{4}$$

it is easy to see that both forecasters expect a positive comparative gain.

Theorem 3. Both forecasters always expect a positive comparative gain for themselves.

As proof, it is clear that

$$\begin{aligned} P_p[CG(X,p,q)] &= P_p[S(X,p)] - P_p[S(X,q)] = \\ &= -H(p) + CH(p||q) - J(p) + CJ(p||q)\,; \\ \text{and similarly,} & \\ P_q[CG(X,q,p)] &= P_q[S(X,q)] - P_q[S(X,p)] = \\ &= -H(q) + CH(q||p) - J(q) + CJ(q||p)\,, \end{aligned} \tag{5}$$

using the four fundamental previsions we identified in Theorem 1. Both of these previsions are positive because the scoring rule is proper.

Now comparing the equation pairs numbered (3) and (5), it is apparent that *summing* each of the paired expected net gains and comparative gains yields precisely the negative results of one another. It is worth stating this too as a theorem.

Theorem 4. $P_p[NG(X,q)] + P_q[NG(X,p)] =$
$$- \{P_p[CG(X,p,q)] + P_q[CG(X,q,p)]\}\ .$$

Notice however that the paired summands for these opposite magnitudes are different.

We shall now continue this analysis by devising an improved procedure for the method of comparative scoring, and then we shall use this result to extend our understanding of Kullback's symmetric divergence.

3. Evaluating pmv's by trading inhering Net Gains

Since p assesses the personal net gain $NG(X,p)$ with prevision 0 while valuing $NG(X,q)$ with the non-zero prevision specified in equation (3), p would surely be willing to trade a claim to plus or minus $NG(X,p)$ in exchange for a claim to $NG(X,q)$ if $P_p[NG(X,q)]$ is positive, and to $-NG(X,q)$ if that prevision is negative. For p assesses $NG(X,p)$ with value 0, whereas either

$NG(X,q)$ or $-NG(X,q)$ is assessed with positive value. In the same way, q values claims to $NG(X,q)$ with prevision 0, while valuing the net gain of p with the prevision $P_q[NG(X,p)]$, as specified in in the lower half of equation (3). Thus both p and q would be pleased to make an exchange in which p receives (appropriately either the positive or negative) value of the $NG(X,q)$ from q while q receives similarly the value of the net gain $NG(X,p)$ (or its negative) from p. In this exchange, both p and q would each be providing the other with something personally regarded as worthless. Thus, both of them would be pleased by the exchanges. Altogether they constitute a "Pareto optimal" exchange.

To clarify this waffle about receiving "a net gain or its negative", note for example that q is avowedly willing to buy *or to sell* the value of $S(X,q)$ for $P_q[S(X,q)]$, since q values these equally. Now with these two options available from q as proclaimed, p would willingly take up q's offer to to sell $S(X,q)$ at this price if $P_p[S(X,q)] > P_q[S(X,q)]$, *or* to take up q's offer to buy $S(X,q)$ at this price if $P_p[S(X,q)] < P_q[S(X,q)]$. The same structure of relative preference applies to q in reacting to p's proclaimed indifference to $S(X,p)$ and $P_p[S(X,p)]$. To simplify the formalities of this waffling about the action whether to buy or to sell, during further discussion we shall denote the content of such exchanges by saying $NG^*(X,q)$ is exchanged for $NG^*(X,p)$. This may mean that p gives $NG(X,p)$ to q in return for $P_p[NG(X,p)]$ *or* that p gives $P_p[NG(X,p)]$ to q in exchange for $NG(X,p)$, at the choice of q. Simultaneously, q gives to p the value of $NG(X,q)$ in return for $P_q[NG(X,q)]$ or vice versa at the choice of p. Of course this specificity would need to be recognized and recorded in the computation of the resulting net gains.

One difference in the assessed values of the two sides of such an exchange is that both p and q would assess their own net gains being given up with a different variance than the net gains they would receive. This constitutes no problem for the usual characterisation of coherent prevision nor for the acceptability of the proposed exchanges. In defining prevision, the scales of exchange are overtly kept small enough that linearity of utility applies to practical valuations. However, we can rescale these agreeable exchanges so that both p and q would regard the two sides of their exchanges symmetrically with the same variance, that is, the same risk. To facilitate this redesign, the pmv proponents would want to rescale the two sides of the exchanges they offer by their assessed standard deviations for them. That is,

p would offer to give q $\frac{NG^*(x,p)}{SD_p[NG(X,p)]}$ in exchange for $\frac{NG^*(x,q)}{SD_p[NG(X,q)]}$;

while at the same time

q would offer to give p $\frac{NG^*(x,q)}{SD_q[NG(X,q)]}$ in exchange for $\frac{NG^*(x,p)}{SD_q[NG(X,p)]}$.

Both p and q regard both of these offers with positive value, so both would be accepted.

Thus, we have found an intriguing and novel solution to the question of how to evaluate the relative quality of p's and q's pmv assertions when the value of X is actually observed. Simplifying the net result of those two exchanges that are agreeable to both of them, the appropriate way to score the pmv assertions of p and q would be

to award to q the score of $\frac{NG^*(x,p)}{SD_p[NG(X,p)]} + \frac{NG^*(x,p)}{SD_q[NG(X,p)]}$, and

to award to p the score of $\frac{NG^*(x,q)}{SD_p[NG(X,q)]} + \frac{NG^*(x,q)}{SD_q[NG(X,q)]}$.

These exchanges constitute for us a *Pareto optimal scoring procedure for comparing the relative quality of p's and q's forecasting distributions*: to award to q the amount of p's achieved Net Gain* scaled by the sum of the square roots of p's and q's precisions for that Net Gain*, and to award to p the amount of q's Net Gain* scaled by the sum of the square roots of p's and q's precisions for *that* Net Gain*.

In order to compute these relative information scores, we need to compute four relative variances: $V_p\{NG[S(X,p)]\}, V_p\{NG[S(X,q)]\}, V_q\{NG[S(X,p)]\}$, and $V_q\{NG[S(X,q)]\}$. Let us digress momentarily to think about these variances in a brief Section of its own.

4. ... and as to the variances

Here are some considerations regarding the cross assessments of the standard deviations. Firstly, remember that $NG(X,p) = S(X,p) - P_p[S(X,p)]$. So $V_p[NG(X,p)] = V_p[S(X,p)]$, because $P_p[S(X,p)]$ is a specified number. Now think of the total log score in the form of $S(X,p) \; = \; log(\frac{p^o}{1-p^o}) \; + \; \sum log(1-p)$ identified in the third line of equation (1). In this form it is clear that

$$V_p S(X,p) = V_p[log(\tfrac{p^o}{1-p^o})] \;\;, \;\; \text{and similarly,} \;\; V_p S(X,q) = V_p[log(\tfrac{q^o}{1-q^o})] \;\; .$$

These follow for the same reason, that the summations $\sum log(1-p)$ and $\sum log(1-q)$ are specified numbers.

Computationally, $\quad V_p[S(X,p)] \; = \; \sum \; p_i \; [log(\frac{p_i}{1-p_i})]^2 \; - \; [\sum \; p_i \; log(\frac{p_i}{1-p_i})]^2 \; ,$

while $\quad V_p[S(X,q)] \; = \; \sum \; p_i \; [log(\frac{q_i}{1-q_i})]^2 \; - \; [\sum \; p_i \; log(\frac{q_i}{1-q_i})]^2$

according to the standard result for any quantity, that $V(Z) = P(Z^2) \; - \; [P(Z)]^2$.

5. Kullback's Symmetric Divergence Complex

In this Section we shall learn how and why the structure of Kullback's symmetric divergence function naturally suggests that it be embellished by vectorial components in three independent dimensions.

To begin, recall the details of the Kullback-Leibler *directed* divergence function $D(\cdot||\cdot)$ and its extropic complement $D^c(\cdot||\cdot)$. These functions are also known as the relative entropy and extropy functions of p with respect to q:

$$\begin{aligned}
D(p||q) \;\; &= \;\; \sum p \; log(\frac{p}{q}) \qquad = \;\; \sum p \; [log(p) - log(q)] \; > \; 0 \; , \text{ and} \\
D^c(p||q) \; &= \; \sum (1-p) \; log(\frac{1-p}{1-q}) \;\; = \;\; \sum (1-p) \; [log(1-p) - log(1-q)] \; > \; 0 \; ;
\end{aligned}$$

and similarly for the "reverse directions",

$$\begin{aligned}
D(q||p) \;\; &= \;\; \sum q \; log(\frac{q}{p}) \qquad = \;\; \sum q \; [log(q) - log(p)] \; > \; 0 \; , \text{ and} \\
D^c(q||p) \; &= \; \sum (1-q) \; log(\frac{1-q}{1-p}) \;\; = \;\; \sum (1-q) \; [log(1-q) - log(1-p)] \; > \; 0 \;\; .
\end{aligned}$$

Those strict exceedance inequalities presume only that the pmvs p, q are not identical: $p \neq q$. Kullback's (1959, p.6) symmetric divergence function $\mathscr{D}(p||q)$ is generated by summing the non-symmetric directed divergences: $\mathscr{D}(p||q) \equiv D(p||q) + D(q||p)$. Similarly for the complementary symmetric divergence, $\mathscr{D}^c(p||q) \equiv D^c(p||q) + D^c(q||p)$. The "total symmetric divergence" is denoted by $\mathscr{D}_T(p||q) = \mathscr{D}(p,q) + \mathscr{D}^c(p,q)$.

5.1. Expected comparative gains sum to the total divergence

Now an intriguing idea derives from examining once again the expected comparative gain equations, this time algebraically in the context of the Total Log Scoring function, and recognising their relations to the Kullback-Leibler *directed divergence* functions. We state this as a

Corollary 1 to Theorem 3. In the context of the Total Log Scoring function $S_{TL}(X, p)$ introduced in equation (1), the previsions for the comparative gain reduce algebraically by direct substitutions to

$$P_p[CG(X,p,q)] = \sum p\, log(\tfrac{p}{q}) + \sum(1-p)\, log(\tfrac{1-p}{1-q}) = D(p||q) + D^c(p||q) \;\; ;$$

and similarly,

$$P_q[CG(X,q,p)] \; = \; \sum q\, log(\tfrac{q}{p}) \; + \; \sum(1-q)\, log(\tfrac{1-q}{1-p}) \; = \; D(q||p) + D^c(q||p) \quad .$$

Each pmv proponent's expected comparative gain over the other's is equal to the relative entropy plus the relative extropy of the proponent's own pmv with respect to the pmv of the other. This is the Kullback-Leibler directed distance, $D(\cdot||\cdot)$, plus its complementary dual, $D^c(\cdot||\cdot)$, both of which are known to be positive. So both the proponents of p and q expect a positive comparative gain in total score relative to the other. The positivity of each's expected gain over the other does not depend on which of them expects to receive a larger score personally. Again, both of them expect to receive a larger score than the other (on account of the propriety of the scoring rule) but surely the size of only one of their own actual scores will be larger than that of the other. In fact, their actual comparative gains are the negative equivalents of one another, according to their defining equation (4).

It is now worth noting that the *sum* of p's and q's previsions for their comparative gains can be seen in equations (5) to equal the *total* symmetric divergence of the pmv's.

Corollary 2 to Theorem 3. Again in the case of the Total Log Score,

$$\begin{aligned} P_p[CG(X,p,q)] + P_q[CG(X,q,p)] &= D(p||q) + D^c(p||q) \ + \ D(q||p) + D^c(q||p) \\ &= \mathscr{D}(p||q) \ + \ \mathscr{D}^c(p||q) \\ &\equiv \mathscr{D}_T(p||q) \ . \end{aligned}$$

This is despite the fact that the sum of their *actual* comparative gains equals 0. Further investigation yields still another intriguing equivalence.

5.2. Expected Net Gains sum to the same thing

Intrigue arises from examining now the algebraic details of p's and q's expectations of each others' *net gains*, viz.,

$$\begin{aligned} P_p\{S(X,q) - P_q[S(X,q)]\} &= \sum(p-q)\ log(q) \ + \ \sum(1-p-(1-q))\ log(1-q) \\ &= \sum(p-q)\ log(\frac{q}{1-q}) \ , \quad \text{and similarly} \\ P_q\{S(X,p) - P_p[S(X,p)]\} &= \sum(q-p)\ log(\frac{p}{1-p}) \ \ . \end{aligned}$$

Each of these is the difference between one's own expection and the other's expectation of the other's log odds for the occuring event.

Now summing *these* expected net gains yields a companion to Corollary 2 of Theorem 3, evident from basic definitions:

Corollary 3 of Theorem 3. Again using the Total Log Scoring rule,

$$P_p[NG(X,q)] \ + \ P_q[NG(X,p)] \ = \ -\mathscr{D}_T(p||q) \ .$$

5.3. An alternative generator of $\mathscr{D}_T(p||q)$.

Perhaps surprisingly, at this late stage in the study of the divergence function, we can now recognise that still two more companion functions can be identified whose sums yield the same symmetric divergence as do the

Kullback-Leibler functions $D(p||q)$ and $D(q||p)$. The awareness arises from Corollaries 2 and 3 of Theorem 3. The first pair, also directed functions, arises formally from switching the components of the finite discrete entropy function, $-H(p) = \sum p \; log \; p$ that get "differenced" in the production of $D(\cdot||\cdot)$. Rather than thinking about $\sum p \; (log \; p - log \; q)$, we consider differencing the "p" instead of the "$log \; p$", raising consideration of $\sum (p-q) \; log \; p$. We shall denote this alternative directed function by $\Delta(\cdot||\cdot)$, motivating it below:

$$\Delta(p||q) \quad \equiv \quad \sum \, (p-q) \; log(p) \quad \text{, and} \tag{6}$$
$$\Delta^c(p||q) \quad \equiv \quad \sum (1-p-(1-q)) \; log(1-p) \; = \; \sum (q-p) \; log(1-p) \quad ;$$

and then similarly for the "reverse directions",

$$\Delta(q||p) \quad = \quad \sum (q-p) \; log(q) \quad \text{, and}$$
$$\Delta^c(q||p) \quad = \quad \sum (p-q) \; log(1-q) \quad .$$

Referring to our formulation of cross prevision assertions, it is apparent that $\Delta(p||q) \; = \; CH(q,p) - H(p)$, the amount by which q's cross entropy for p exceeds the entropy in p itself. Similarly, $\Delta^c(q||p) \; = \; CJ(q,p) - J(p)$, the amount by which q's cross extropy for q exceeds q's own extropy.

These considerations yield a surprising result: both the directed distance $D(p||q)$ and its reverse $D(q||p)$, and this new function $\Delta(p||q)$ and its reverse $\Delta(q||p)$, *sum* to the same thing, ... Kullback's symmetric divergence, which we are denoting by $\mathscr{D}(p||q)$. The same is true of the complementary sums, which we denote by $\mathscr{D}^c(p||q)$. This result arises as

Corollary 1 to Theorem 4.

$$\begin{aligned} \Delta(p||q) \, + \, \Delta(q||p) \; &= \; \Sigma(p-q) \; [log(p) - log(q)] \; = \\ &= \; D(p||q) \, + \, D(q||p) = \; \mathscr{D}(p||q), \end{aligned}$$

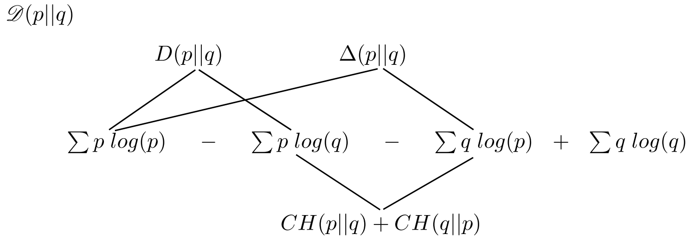
and

$$\begin{aligned} \Delta^c(p||q) + \Delta^c(q||p) \; &= \; \Sigma(q-p)[log(1-p) - log(1-q)] \; = \\ &= \; D^c(p||q) + D^c(q||p) \; = \; \mathscr{D}^c(p||q). \end{aligned}$$

Here the sums of the $\Delta(\cdot||\cdot)$ functions apply the total log score to the Net Gains, while the $D(\cdot||\cdot)$ functions arise from its application to the Comparative Gains.

5.4. Identifying still a third function that yields $\mathscr{D}_T(p,q)$

Expanding the identical sums $D(p||q)+D(q||p)$ and $\Delta(p||q)+\Delta(q||p)$ which constitute $\mathscr{D}(p||q)$ into their component pieces, and then exhibiting them across a line makes it apparent that each of the functions, $D(p||q)$ and $\Delta(p||q)$ arises by selecting two of the four summands to define the function, and thus the remaining two to define their "reverses", $D(q||p)$ and $\Delta(q||p)$.



$D(p||q)$ is the sum of the first two pieces of the first line, leaving $D(q||p)$ composed of the second two; while $\Delta(p||q)$ is the sum of the first and third piece, leaving $\Delta(q||p)$ to be composed of the second and fourth piece.

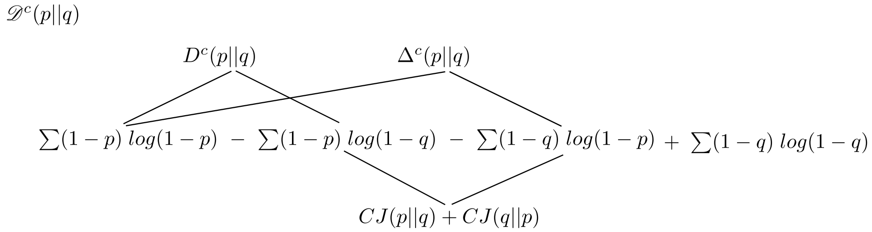
Now seeing these sums and these function pairs in this way suggests that a third (and the accompanying sixth) choice of two of the four summands may be worthy of study as well. Consider using the interior two summands (the second and third pieces of the first line) to constitute still another form of function that would sum (along with the remaining pair) to yield the symmetric divergence $\mathscr{D}(p||q)$. This choice would be the sum of the "cross entropy" functions, $CH(p||q)+CH(q||p)$. One difference of this function from $D(p||q)$ and $\Delta(p||q)$ is that this sum of cross entropies is already symmetric on its own, whereas the latter two functions require summation with their companions $D(q||p)$ and $\Delta(q||p)$ to construct a symmetric function. A second difference of the sum $CH(p||q)+CH(q||p)$ from the $D(p||q)$ and $\Delta(p||q)$ functions appears in its location for self-divergence. Whereas the values $D(p||p) = \Delta(p||p) = 0$, the value of $CH(p||p)+CH(q||q) = H(p)+H(q)$, the summed entropies in p and in q. However, defining the symmetric divergence function more generally than heretofore so to eliminate this location shift of "self-distance" yields a pleasing symmetrical result.

Theorem 5. Based on a generalised definition of symmetric divergence, so to account for the location of "self-divergence",

$$\begin{aligned}
\mathscr{D}(p||q) &\equiv D(p||q) + D(q||p) - D(p||p) - D(q||q) \\
&= \Delta(p||q) + \Delta(q||p) - \Delta(p||p) - \Delta(q||q) \\
&= CH(p||q) + CH(q||p) - CH(p||p) - CH(q||q) \\
&\geq 0\,.
\end{aligned} \tag{7}$$

Of course those self-divergence values in the first two lines equal 0, whereas in the third line $CH(p||p) = H(p)$ and $CH(q||q) = H(q)$.

The same procedures of deconstruction and reconstruction can be performed with the complementary functions $(D^c(\cdot||\cdot)$ and $\Delta^c(\cdot||\cdot))$ in a second line: $D^c(p||q)$ is the sum of the first two pieces of this second line, while $\Delta^c(p||q)$ is the sum of the first and third pieces.



We propose that the Kullback directed distance function needs to be appended by all three of its three generators if it is to characterise the information content of the statistical problem the forecasters address. Together they constitute a complex.

One might be puzzled. "Who cares about all these generators of the same symmetric distance? The symmetric distance is what it is." The answer derives from an identifiable linear relation between the four functions $\mathscr{D}_T(p||q)$, $D(p||q) + D^c(p||q)$, $\Delta(p||q) + \Delta^c(p||q)$, $\mathscr{C}_H(p||q) + \mathscr{C}_J(p||q)$, where $\mathscr{C}_H(p||q) \equiv CH(p||q) + CH(q||p)$ and $\mathscr{C}_J(p||q) \equiv CJ(p||q) + CJ(q||p)$,[2] and the "four fundamental previsions" we introduced in equation (2). We address this now.

5.5. The isomorphism of the Kullback complex of generators with the four fundamental previsions

Each of these eight component functions in a Kullback complex is a different linear combination of the various entropies, extropies, cross entropies and cross extropies that have been central to our scoring analysis. These combinations happen to be ordered in such a way that the two spaces spanned by the 4-dimensional vector functions are isomorphic, related by the linear equations

$$\begin{bmatrix} \mathscr{D}(p||q)+\mathscr{D}^c(p||q) \\ D(p||q)+D^c(p||q) \\ \Delta(p||q)+\Delta^c(p||q) \\ \mathscr{C}_H(p||q)+\mathscr{C}_J(p||q) \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & -1 \end{bmatrix} \times \begin{bmatrix} P_p[S(X,p)] \\ P_p[S(X,q)] \\ P_q[S(X,q)] \\ P_q[S(X,p)] \end{bmatrix}. \quad (8)$$

The inverse transformation is

$$\begin{bmatrix} P_p[S(X,p)] \\ P_p[S(X,q)] \\ P_q[S(X,q)] \\ P_q[S(X,p)] \end{bmatrix} = \begin{bmatrix} 0 & .5 & .5 & -.5 \\ 0 & -.5 & .5 & -.5 \\ 1 & -.5 & -.5 & -.5 \\ 0 & .5 & -.5 & -.5 \end{bmatrix} \times \begin{bmatrix} \mathscr{D}(p||q)+\mathscr{D}^c(p||q) \\ D(p||q)+D^c(p||q) \\ \Delta(p||q)+\Delta^c(p||q) \\ \mathscr{C}_H(p||q)+\mathscr{C}_J(p||q) \end{bmatrix}. \quad (9)$$

What this relation tells us is that each of the three distance function generating pairs measures a distinct information source for understanding of the full information content of Kullback's symmetric distance. The symmetric distance measure is incomplete on its own: it needs to be supplemented by three more components if it is to represent the information content of the pmv assertions p and q. In all, it is identifiable better as a *vector* which we might refer to as the Kullback complex ... $[\mathscr{D}_T(p||q),\ D(p||q)+D^c(p||q),\ \Delta(p||q)+\Delta^c(p||q), \mathscr{C}_H(p||q)+\mathscr{C}_J(p||q)]$, rather than merely by the symmetric function value $\mathscr{D}_T(p||q)$ alone.

One further recognition is in order. We have couched this discussion in the context of the total logarithmic score (1) and the Kullback total symmetric distance function $\mathscr{D}_T(p||q)$. It is worth remarking that the total log score can be considered to be the sum of two separate components: the log score itself, $\sum E_i\, log\, p_i$, and the complementary log score, $\sum \tilde{E}_i\, log\,(1-p_i)$. Actually, both of these components of the total log score are proper scores on their own, as is *any* linear combination of them. The two component scores assess different aspects of the asserted pmv's p and q. One piece focuses more directly on the center of the mass function, while the other focuses more on the tails. Because these aspects of a pmv may concern different people in different ways with respect to utility, it is usually best to keep track of the two components of the score separately. If the scoring function $S_{TL}(X, \mathbf{p}_N)$ were considered to be the sum of the two separate scores, say, $S_{log}(X, \mathbf{p}_N)$ + $S_{complog}(X, \mathbf{p}_N)$, the isomorphism equations (9) hold separately for each component of the total log score and their previsions too. Tracking these components separately, the total symmetric Kullback complex could be considered to be an 8-dimensional space of measurement functions.

We shall now address a final technical note concerning the relation of these newly recognised dimensions of the information complex to the theory of Bregman divergences.

6. $\Delta(\cdot||\cdot)$ and $HC(\cdot||\cdot)$ are *not* Bregman divergences

Information theorists are comfortable by now with the fact the Kullback's symmetric divergence function is the sum of two "directed divergences", each summand of which is a Bregman divergence. That is, $\mathscr{D}(p||q) = D(p||q) + D(q||p)$, and $D(p||q)$ is a Bregman divergence with respect to the Bregman function $\Phi(p) = \sum p\ log\ p$. The same is true of the complementary Kullback symmetric divergence: $\mathscr{D}^c(p||q) = D^c(p||q) + D^c(q||p)$, where $D^c(p||q)$ is the Bregman divergence associated with the Bregman function $\Phi^c(p) = \sum(1-p)\ log\ (1-p)$. We shall remind you of the definition of a Bregman divergence shortly.

It may come as a surprise then that *neither* of the companion generators of Kullback's symmetric function that we have found constitutes a Bregman divergence. These are the functions we have denoted by $\Delta(p||q) \equiv \sum(p-q)\ log\ p$, and $CH(p||q) \equiv \sum p\ log\ q$. In brief, a function $d_\Phi(p,q)$ is a Bregman divergence with respect to a convex, differentiable function $\Phi(\mathbf{p}_N)$ if $d_\Phi(p,q) = \sum [\phi(p) - \phi(q) - (p-q)\phi'(q)]$. To be precise, this describes only the "separable" context in which $\Phi(\mathbf{p}_N)$ has the form

$\sum \phi(p_i)$, where $\phi(\cdot)$ is a function of the single component variables of the vector p. However, this *is* the limited context that is relevant to our discussion here.

Using some simple algebraic adjustments then, $\Delta(p||q)$ and $CH(p||q)$ would also be identifiable as Bregman functions if and only if there were solutions $\phi(\cdot)$ and $\gamma(\cdot)$ to the following two problems of differential equations:

1. Find a convex differentiable function $\phi(\cdot)$, with a domain that includes both p and q in the interval $(0,1)$, and for which

$$\sum (p-q)\, log\ p \ = \ \sum [\phi(p) - \phi(q) - (p-q)\phi'(q)]$$

or equivalently for which, $\phi'(q) \ = \ \frac{\phi(p)-\phi(q)}{p-q} \ - \ log\ p$ for any p and $q \ \in \ (0,1)$; and

2. Find another such convex differentiable function $\gamma(\cdot)$ for which

$$\sum p\ log\ q \ = \ \sum [\gamma(p) - \gamma(q) - (p-q)\gamma'(q)]$$

or equivalently for which, $\gamma'(q) \ = \ \frac{\gamma(p)-\gamma(q)}{p-q} \ - \ \frac{p\ log\ q}{p-q}$ for any p and $q \ \in \ (0,1)$.

If such functions could be found, then they would be the Bregman functions that specify $\Delta(p||q)$ and $CH(p||q)$ as their Bregman divergences, respectively.

However, neither of these problems does have solutions! This is revealed by examining the limits as $p \rightarrow q$ of these two conditioning equations on the sought-for functions. In this limiting context, the first problem would require that $\phi'(q) \ = \ \phi'(q) + log\ q$; while the second would require another nonsense ... that $\phi'(q) \ = \ \phi'(q) + \infty$. On the contrary, if the left-hand-side of the conditions of problems 1 or 2 were replaced by Kullback-Leibler's function $D(p||q)$, then this limiting condition would require only that $\phi'(q) \ = \ \phi'(q) + lim_{p \rightarrow q} \frac{p\ log\ p - p\ log\ q}{p-q} \ - \ 1$, a condition that is assured by L'Hopital's rule. This feature is what distinguishes the Kullback-Leibler measure as a Bregman function.

The upshot of this understanding is that the dissection of of $\mathscr{D}(p,q)$ via each of the three directed functions $D(\cdot||\cdot), \Delta(\cdot||\cdot)$, and $CH(\cdot||\cdot)$ illuminates a distinct dimension of the measure we know as Kullback's total symmetric distance. We might well call the complete vector of these measures a "Kullback complex." All arguments regarding the accompanying dimensions of $\mathscr{D}_T(p||q)$ pertain individually to its direct and complementary components, $\mathscr{D}(p||q)$ and $\mathscr{D}^c(p||q)$, as well.

A concluding query arises then from this analysis. What is the real motivating feature of Bregman divergences that makes them so widely well regarded? We have found out that two other equally fundamental measures generate the same symmetric divergence as does the Kullback divergence, but these newly recognised generating functions are *not* Bregman divergences. Evidently they do provide substantive content to a complex of measures which join with the symmetric divergence itself to summarize information.

7. The subjectivist take on the Kullback functions

Kullback's classic work (1959) on *Information Theory and Statistics* developed out of quite a different statistical imagination than the subjectivist construction we have been developing here. He began (pp. 4-5) by positing two hypothesized random generating functions of an observation, X, described by the densities $f_1(x)$ and $f_2(x)$. Ignoring here some measure theoretic refinements, he then identified the measure $\int f_1(x)\ log(\frac{f_1(x)}{f_2(x)})\ dx \ \equiv \ I(1:2)$ as "the information in $X = x$ for discriminating in favour of the generator $f_1(x)$ as opposed to $f_2(x)$." This choice was motivated by the measure's role in distinguishing the expected posterior log odds from the prior log odds in favour of H_1 relative to H_2. It was this nonsymmetric "information function" $I(1:2)$ which he joined with its reverse function $I(2:1)$ to yield his symmetric divergence function $J(1,2) = I(1:2) + I(2:1)$.

Indeed, this framework currently still provides the standard objectivist context for most all contemporary work in information statistics. The subjectivist context in which we honour the pathbreaking insight of Bruno de Finetti construes all statistical problems quite differently.

In the subjectivist context of the work we have been addressing here, the notation for Kullback's symmetric divergence function $J(1,2)$ has been replaced notationally by $\mathscr{D}(p,q)$, and is used for another purpose. The functions $f_1(x)$ and $f_2(x)$ are replaced by our vectors p and q. They are understood to represent not "unobservable potential random generators of X" that need to be "tested", but as pmv vectors of personally asserted probabilities for the possible values of X by two different people based on their differing information sources and assessments. Kullback's non-symmetric information function $I(1:2)$ has been denoted herein as $D(p||q)$. We now find that two other functions, $\Delta(p||q)$ and $CH(p||q)$, can be used to generate the same symmetric divergence function $\mathscr{D}(p||q)$ as does $D(p||q)$ when it is defined more completely, and we know why. The Kullback complex is merely a translation of the information measures encoded by our four fundamental expectation equations into different ranges and scales. The information components $D(\cdot||\cdot)$, $\Delta(\cdot||\cdot)$, and $CH(\cdot||\cdot)$ are linearly distinct from $\mathscr{D}(\cdot||\cdot)$ itself.

8. On statistical application

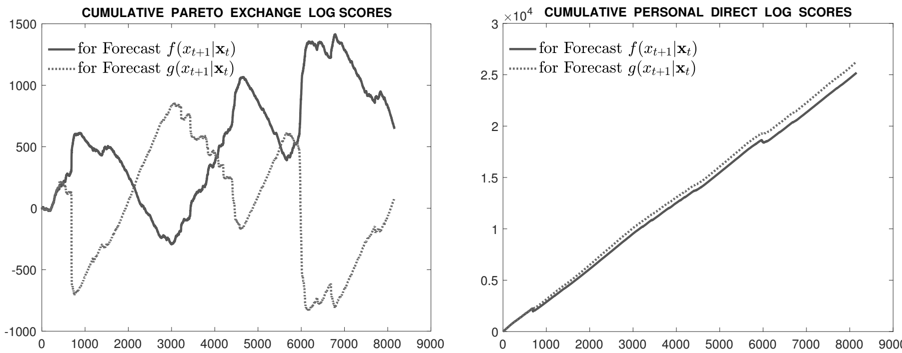
We have appreciated the opportunity to present this analysis as a tribute to the marvelous insights provided by the work of Bruno de Finetti. It is only the limitation of space that preclude a complete exhibition of a substantive application of its implications here. Two remarks are in order.

8.1. A preliminary report

The authors have been engaged for several years in various applications of the comparative evaluation of empirical forecasting distributions using proper scoring rules. In so doing we have used the the traditional style of awarding forecasters the proper score appropriate to their own forecasts. However, we had been bothered by the arbitrary aspects involved in that method that we have mentioned in Section 2.2. In collaboration with our colleague Gianna Agrò, we are currently engaged in a reassessment of a proper scoring analysis of alternative forecast distributions for daily stock returns that exhibit strikingly different structures in the tail properties of the forecasting distributions. Our original analysis conducted in the traditional style was reported in Agrò et al. (2010). Though our work on the reassessment is yet incomplete, we are already aware that the implications of the Pareto exchange methodology make major differences in the assessment of the two forecast distributions. An analysis of the complete results of this research will soon be accessible. It is currently only in working form in Sanfilippo et al. (2017).

To provide a taste of the strength of the difference the Pareto exchange makes in applications, we present here in Figure 1 a side-by-side comparison of the cumulative Pareto-exchange scores and the simple cumulative personal scores for two mass functions. Without further description here, they are identified only as Forecast density $f(x_{t+1}|\mathbf{x}_t)$ and Forecast density $g(x_{t+1}|\mathbf{x}_t)$. The differences in the historical structures of the scoring behaviour are startling. Rather than suggesting a regular improvement in the forecast of "g" upon that of "f" as seen in Figure 1b according to Personal Direct Log Scoring, the Cumulative Pareto Exchange Log Score displays changing periods of dominance by each of the alternative forecasts in Figure 1a. Moreover, these are left in a cumulative state that favours the Forecast "f" rather than "g". A complete specification of precise details must be delayed to the full report.

Figure 1. **Comparative results of Pareto exchange Scores and Direct Scores**



(a) The Pareto exchange of normed Log Scores identifies periodic shifts in the assessed quality of forecast performances by "f" and "g".

(b) The direct accumulation of Log Scores suggests regularly improved quality of forecast "g" relative to that of "f".

8.2. The subjectivist alternative to hypothesis testing

In our introductory remarks we mentioned that from de Finetti's subjectivist perspective, the time honoured practice of statistical hypothesis testing needs to be replaced by methods based on proper scoring rules. Without presenting an exemplary analysis, we can sketch here the structure of such a procedure based on the Pareto optimal exchange of proper scores.

Consider the context of the simplest statistical problem for which the Neyman-Pearson Lemma applies. A sequence of quantities presumed to be independent emanations of a distribution with mean μ are to be observed, and it is proposed to test the null hypothesis H_0 that $\mu = \mu_0$ against an alternative H_A that $\mu = \mu_A$. Despite early recognition in Lehmann (1959, p 61; p 69 in Second Edition) that the choice of an appropriate "significance level" required for performing such a test needs be based on a decision maker's relative utility for the correct and erroneous choice of the two hypotheses, statistical practice has drifted toward the understanding of arbitrary significance levels $.05, .01$, and .001 as summary guides to an appropriate decision to "reject" or to "accept" the null hypothesis. More egregiously, the summary has been long extended to the computation and report of a "p-value".

Bruno de Finetti's subjectivist outlook would characterise this situation quite differently. The sequence of observations are recognised as meriting a judgment of exchangeability, represented by a distribution which is necessarily then mixture-hypergeometric. The relative forecast quality of two such distributions could then be evaluated by methods exposed here. One would be based upon a mixture favouring the location μ_0, and the other favouring μ_A. The accumulating sequential scores of their forecasts would involve a Pareto exchange of their proper scores. An exposition of a simple computational example in this standard context would be appropriate to introductory level texts on applied statistics.

9. Concluding remark

A little more than 50 years have passed since Buno de Finetti voiced his "partially baked" musing in I. J. Good's anthology: "Does it make sense to speak of 'good probability appraisers'?" He had wondered about this question at that time because of the subjectivist attitude, which we support, that no probability assertion can be determined to be "wrong" on the basis of observational evidence. For the probability assertion merely constitutes an honest expression of its promoter's uncertain knowledge about the value of the unknown quantity. There is nothing wrong with being uncertain. That is our common state. Nonetheless, the analysis we have provided in this tribute essay supports an answer of "Yes!" to de Finetti's topical question. While details of our analysis pertain to the context of the total log score, our identification of the Pareto exchange is applicable to any proper scoring function. Proper scoring rules can be used meaningfully to aid in the evaluation of the information quality in personal probability assertions as well as in elicitation of personal probabilities themselves. However, the harvest of this usefulness requires both a switch in the scores that have commonly been awarded to each of the mass function proponents and an expansion of our understanding of the symmetric Kullback divergence measure. This has long been considered to be the cornerstone standard valuations of statistical information.

Acknowledgements

Thanks to Rua Murray for a suggestion regarding the Bregman derivatives, and to Gianna Agrò for her permission to print here the preliminary Figure 1 which represents the result of our joint work in progress. Thanks to Anuj Misra for producing the visual schema representing the construction of Kullback's symmetric divergences $D(p,q)$ and $D^c(p,q)$.

References

Agr`o, G., Lad, F., and Sanfilippo, G. (2010) *Sequentially forecasting economic indices using mixture linear combinations of EP distributions*, Journal of Data Science, 8, 101-126.

Dawid, A.P. and Musio, M. (2015) *Bayesian model selection based on proper scoring rules*, Bayesian Analysis, 10, 479-499.

de Finetti, B. (1962) *Does it make sense to speak of "good probability apprais-ers"?*, in The Scientist Speculates: an anthology of partly-baked ideas, I. J. Good (ed.), London: William Heinemann, 357-364.

de Finetti, B. (1967) *Quelque conventions qui semblent utiles*, Revue Romaines des Mathe ´matiques Pures et Appliques´, 12, 1227-1233. L.J. Savage (tr.) A useful notation, in Probability, Statistics and Induction: the art of guessing, New York: John Wiley (1972), pp. xvii-xxiv.

Gneiting, T. and Raftery, A. (2007) *Strictly proper scoring rules, prediction, and estimation*, JASA, 102, 359-378.

Good, I.J.. (1962) *The Scientist Speculates: an anthology of partly-baked ideas, London*: William Heinemann. This includes the article of de Finetti (1962).

Kullback, S. (1959) *Information Theory and Statistics*, New York: John Wiley.

Kullback, S. and Leibler, R.A. (1951) *On information and sufficiency*, Annals of Mathematical Statistics, 22, 79-86.

Lad, F. (1996) *Operational Subjective Statistical Methods: a mathematical, philosophical, and historical introduction*, New York: John Wiley.

Lad, F., Sanfilippo, G., and Agr`o, G. (2012) *Completing the logarithmic scor-ing rule for assessing probability distributions*, ISBrA Conference, AIP Conf. Proc. 1490, 13, 13-30; View online: http://dx.doi.org/10.1063/1.4759585

Lad, F., Sanfilippo, G., and Agr`o, G. (2015) *Extropy: complementary dual of Entropy*, Statistical Science, 30, 40-58.

Lehmann, E.L. (1959) T*esting Statistical Hypotheses*, New York: John Wiley.

Matheson, J. and Winkler, R. (1976) *Scoring rules for continuous probability distributions*, Management Science, 22, 1087-1096.

Sanfilippo, G., Lad, F., and Agr`o, G. (2017) *Scoring alternative forecast distri-butions using Pareto exchanges of proper scores*. Working paper.

Savage, L.J. (1971) *Elicitation of personal probabilities and expectations*, JASA, 66, 778-801.

G. & L. E. R., Vol. 22 No. 2, 2018

Francesco Carlucci *

INTERNATIONAL MONETARY SHOCKS AND CURRENCY MOVEMENTS

Abstract

A vector autoregressive (VAR) model linking the economies of the main countries (or monetary zones) in the world – the USA, the Eurozone, Japan and China – is built to analyse the effects on their economies produced by the monetary policy of each one. The economies are characterized by four variables: nominal exchange rate against the US dollar, consumer price, quantity of money in terms of M1, and GDP. The estimation of the VAR model is of the Bayesian type, in order to take into account not only the data but also the opinions of the model builder. Estimation is performed by use of the method of dummy observations as the number of variables is high, twice the maximum normally used in VAR models. A monetary shock in the USA produces the usual effects on economies of the USA themselves, Japan, and the Eurozone, whilst one in the latter increases the output not only in itself but also in the USA. Monetary shocks in Japan and China have poor effects on their own economies.

JEL CLASSIFICATION: C11; C54; E52; F47;
KEYWORDS: Var Model; Bayesian Estimation; Impulse Responses; International Monetary System; Monetary Shock; Exchange Rate

* University of Rome “Sapienza”, via del Castro Laurenziano, n. 9, Rome, Italy. Phone: +39-06-49766354. *E-mail address*: francesco.carlucci@uniroma1.it.

1. Introduction

In response to the great crisis of 2008, the world leading monetary institutions (the Federal Bank in the USA, the European Central Bank, the Central Bank of Japan) first reduced their reference interest rates to nearly zero[1], and then increased their balance sheets through the purchase of assets (quantitative easing) in order to promote consumption and domestic investments. But it can be assumed that these institutions, while not openly acknowledging it, have sought to promote economic activity also through the exchange-rate depreciation, factual consequence of monetary easing[2].

In fact, the exchange-rate dynamics is generally regarded as a by-product of monetary policy, as the exchange rate is rarely, in advanced economies, a specific objective of economic policy. In the long run, the benefits from the exchange-rate depreciation should be close to zero, mainly because of higher prices, while in the short run they might exist, with an improvement of the balance of payments at the expense of balances in competing countries.

These argumentations, however, can be questioned. As for the long run, for example, China has long been accused by the US Congress of "manipulating" the exchange rate of its currency, in order to keep it low and thus promote the domestic economic activity[3]. As for the short run, again by example, the lowering of the exchange rate by a country could make the central banks of the competing countries react by implementing measures leading to a depreciation of their currencies, not accepting that the improvement of a competing economy can occur penalizing their own.

This may be possible. But it may be also possible that, again in the short run, the improvement of the balance of payments of a country, due to the

[1] See Eggertsson and Woodford (2003).
[2] See, for instance, Eichenbaum and Evans (1995).
[3] China could reach this result through a high reserve accumulation and a certain degree of capital controls. See, for instance, Rodrik (2008), who analyses the relationship between reserve accumulation and the real exchange rate, or Korineck (2013), who shows how the existence of a negative externality for the aggregate demand, as an insufficient demand, should lead to the implementation of monetary-expansion policies which depreciate the currency.

increase of money supply, results in an increase of its welfare and thus of imports from competing countries[4].

From this framework, complex and uncertain, it is unfeasible to construct a structural model for the world economic system that makes it possible to analyse the effects of monetary and currency policies of the major countries (or monetary zone with a single currency). Too many equations might be specified in one completely different way or another. Thus, in this paper, where such effects are investigated, the main variables that interact in the system (the currency of a country, its exchange rate, output, price), are linearly interconnected with each other without hypothesizing specific relationships. They are inserted into a vector autoregressive (VAR) model, which determines the effects caused on them by specific impulses of some of them considered as instruments. Vector autoregressions are now standard tools for structural analysis in macroeconomics, as they do not impose restrictions on its parameters and allow to capture complex data relations. The size of the VAR model generally ranges from five to nine variables and so makes it difficult to be used in international economics. Thus in this paper a particular shrinkage methodology has been applied to handle fifteen variables.

We build a VAR model that links the economies of the major countries (or monetary zones) in the world - the United States, the Eurozone, Japan and (the People's Republic of) China - to study the effects caused by the monetary policy of each of them. Their economic systems are characterized by the following four variables: nominal exchange rate against the US dollar, consumer price measured with the exception of energy products and food, amount of money in terms of M1[5], and GDP, which is considered to be an indicator of economic growth[6].

---

[4] More than thirty years ago Eichengreen and Sachs (1985 and 1986) showed that the devaluations in the thirties had not only not led to the collapse of the international monetary system but also contributed to the world's economic recovery.

[5] M1 was preferred to monetary base to better capture the influence of money on economic activity as, for instance, Leeper, Sims, Zha, Hall and Bernanke (1996) do. Christiano, Eichenbaum and Evans (1999), on the contrary, use M2.

[6] It would be interesting to consider also the inflation expectations, but it was not possible owing to the non-availability of data.

In fact, it would have been more correct to take the reference interest rate[7] instead of the quantity of money as the main instrument of monetary policy, but the latter has replaced the former since 2008 and the use of money in the entire sample period, from 1991 through 2016, has been preferred, assuming, for the years before 2008, an approximate (inverse) relationship between them. In this way, monetary policy is represented by only one variable instead of two.

This policy and the exchange-rate one are supposed to be approximately linked, although they operate in a context of imperfect mobility of international capitals. Thus the VAR model allows you to analyse not only the effects of monetary impulses on economic growth and prices, but also how the exchange rate mediates between these impulses and economic activity, of course regardless of the tools used to implement policies.

The method of estimating the model is not classic but Bayesian[8]. This allowed the subjectivist opinions of the model builder to be taken into account besides the available data. It is a relevant opportunity because the model can be shaped by means of subjective views on international economic policies that are not related to statutory commitments or customary behaviours.

In the methodology, each variables is usually assumed *a priori* not to influence the others, but only itself at lag one. In this paper, on the contrary, some corrections have been made to the method to let the variables *a priori* interact with each other (Section 4).

---

[7] As, for instance, Eichenbaum and Evans (1995) do.

[8] The Bayesian Econometrics, which in half a century has been making giant strides and which is used in this paper, was founded by Arnold Zellner, who presented his seminal book "Introduction to Bayesian Econometrics" in a seminar organized in the mid-sixties of the past century by Bruno de Finetti, radical Bayesian. In those years at the Mathematics Department of La Sapienza University, where de Finetti had moved from the Economics Faculty because of health problems, his seminars on Probability, Econometrics, and Mathematical Economics, were frequent. Among others, Koopmans, Lindley, Scarf, Frisch, came to hold the seminars, but there were few participants. There were G. Pala, whom de Finetti had brought as his assistent from the Faculty of Economics, and G. Majone, convinced to leave the USA to teach Mathematical Statistics in Rome (with little luck, because Majone soon returned to the USA). Often came D. Tosato, then young economist, and one or two of de Finetti's undergraduate students.

The following Section sets up the VAR model and lists the variables that compose it. Since the subjective opinions of the model builder must be taken into account, the estimation procedure is Bayesian. But the standard Bayesian estimation method cannot be used because the variables are too many. So, the estimation procedure of Banbura, Giannone and Reichlin (2010) is used as it allows you to handle even several dozens of variables. Section 3 describes such a method while Section 4 expounds the analytical transformation procedure of the initial subjective opinions of the model builder. Sections 5 and 6 report the analyses of the effects of the monetary and exchange-rate policies in the four countries (or monetary zones) performed by two typical analytical techniques: the impulse response functions (Section 5) and the variance decomposition of forecast errors (Section 6). The final Section presents the conclusions.

2. The VAR model and some problems of estimation

Please, So, four variables were considered: the quantity of money in terms of M1[9], the exchange rate national currency/US dollar[10], an index of consumer prices excluding energy products and food[11], the real gross domestic product[12]. For four countries or monetary areas: Japan, the United States, the Eurozone, and (the People's Republic of) China. Because the exchange rate is considered against the dollar, currency of the USA, the exchange rate for this monetary zone has not been used, so that the total number of variables, set in the vector $\mathbf{z}_t$, is 15.

The data set consists of quarterly observations (1991:1-2016:3), which are taken from the OECD data-base of fall 2016. Variables are logarithmized and linked by a VAR model of order 2, as the usual AIC and BIC criteria have suggested the order 1 or 2. Thus, the model is:

$$\mathrm{z_t} = \mathbf{\Phi}_1 \mathrm{z_{t-1}} + \mathbf{\Phi}_2 \mathrm{z_{t-2}} + \mathbf{k} + \mathbf{u}_\mathrm{t} \qquad (1)$$

---

[9] Index numbers, 100 in 2010. Eurozone with 19 countries.
[10] Eurozone with 19 countries.
[11] Index numbers, 100 in 2005 for Japan, in 2009 for the USA, in 2005 for China. Eurozone with 19 countries.
[12] Market prices.

where $\mathbf{\Phi_1}$ and $\mathbf{\Phi_2}$ are square matrices of order 15 and $\mathbf{k} = [k_1, k_2, \dots, k_{15}]'$ is a vector of 15 constants; is $\tilde{\mathbf{u}}_t$ supposed to be a vector white noise, with dispersion matrix $\sum_u = E(\tilde{\mathbf{u}}_t \cdot \tilde{\mathbf{u}}_t{}')$[13].

The classical estimation of the parameters of the model (1) does not present any difficulty, since the ordinary least squares (OLS) can be used equation by equation, after Cholesky factorizing Σ_u. The Bayesian one, on the contrary, necessary to take account of the subjective opinions of the model builder, is difficult, not to say impossible. Indeed, if you want to include these elements even in the simplest *a priori* distribution (which allows one to take account of the subjective opinions), that by Litterman (1986), also known as Minnesota, when estimating parameters and calculating the impulse response functions, some square matrices of order $15 \cdot (15 \cdot 2 + 1) = 465$ must be inverted. This inversion is computationally not simple, due to the number of variables.

On the contrary, the estimation of parameters of (1) by means of ordinary least squares (OLS) and the calculation of impulse response functions require the inversion of square matrices of order $k = 15 \cdot 2 + 1 = 31$, computationally manageable in a simple way.

But there is a procedure, shown by Banbura, Giannone and Reichlin (2010), which allows estimating the model (1) by OLS and simultaneously utilizing subjective elements in an *a priori* distribution similar to the Minnesota one. It uses additional dummy observations. Let's see it, as it will be used for the objectives of the paper.

3. The Bayesian estimation

The procedure uses the OLS estimation, equation by equation, for the model parameters, assuming that Σ_u is Cholesky factorized[14]. Furthermore, it makes use of the addition of a set of dummy observations, by means of which an *a priori* distribution of the Minnesota type is imposed. The 15 variables are inserted according to the following ordering: M1-USA (US money), M1-EUR (Eurozone money), EXC-EUR (euro/US dollar exchange rate), M1-JAP (Japanese currency), GDP-USA (US GDP), GDP-EUR, M1-CHI (China money), EXC-JAP, PRI-USA (US price), PRI-JAP, EXC-CHI,

[13] A tilde over its symbol points out that the variable is considered in its stochastic version.

[14] The Cholesky factorization in the VAR analysis was firstly used by Sims (1980).

PRI-EUR, GDP-JAP, GDP-CHI, PRI-CHI. This ordering is defined to set (approximately) earlier the variables that have the greatest impact on the others, and at the end those considered less significant in explaining the others.

The VAR(2) model given by (1) may be written, for *t*=1, 2,…,*n*, in the usual regressive form:

$$\mathbf{Z}=\mathbf{X}\mathbf{B}+\mathbf{U}$$

where, $\underset{(n\times 15)}{\mathbf{Z}} = [\boldsymbol{z}_1\ \boldsymbol{z}_2 \dots \dots \boldsymbol{z}_n]'$, $\underset{(n\times 15)}{\mathbf{U}} = [\boldsymbol{u}_1\ \boldsymbol{u}_2 \dots \dots \boldsymbol{u}_n]'$, $\underset{(\mathrm{n}\times \mathrm{k})}{\mathbf{X}} = [\mathbf{X}_1\ \mathbf{X}_2 \dots \dots \mathbf{X}_n]'$ with $\underset{k\times 1}{\mathbf{X}_\mathrm{t}} = [\mathbf{z}_{t-1}{}'\ \boldsymbol{z}_{t-2}{}'\ 1]'$, and:

$$\underset{(k\times 15)}{\mathbf{B}} = \begin{bmatrix} \mathbf{\Phi}_1' \\ \boldsymbol{\Phi}_2' \\ k' \end{bmatrix}$$

is the matrix of parameters to estimate. As the variables are numerous, the procedure by Banbura, Giannone and Reichlin (2010) is followed and 33 dummy observations are added to the original data, i.e. the actual observation matrix Z is enlarged by setting over it the following dummy-observation matrix:

$$\underset{(46\times 15)}{\mathbf{Z}_d} = \begin{bmatrix} diag\left\langle \delta_{1,1}^{(1)}\sigma_1, \dots, \delta_{15,15}^{(1)}\sigma_{15} \right\rangle / \lambda \\ \underset{15\cdot(2-1)\times 15}{\mathbf{0}} \\ diag\left\langle \sigma_1, \sigma_2, \dots, \sigma_{15} \right\rangle \\ \underset{1\times 15}{\mathbf{0}} \end{bmatrix}$$

where $\sigma_1, \sigma_2, \dots, \sigma_{15}$ are the standard deviations of disturbances in univariate autoregressive models constructed for the variables, and $\delta_{1,1}^{(1)}, \delta_{2,2}^{(1)}, \dots, \delta_{15,15}^{(1)}$ are the values for the autoregression parameters of variables lagged once in the Minnesota *a priori* distribution. Here the assumption is made that the trend of a variable, if exists, is stochastic and the corresponding $\delta^{(1)}$ is 1. If the trend does not exist and the variable is assumed to be stationary, the corresponding $\delta^{(1)}$ is less than 1 in absolute value[15].

The (shrinkage) hyperparameter λ points out how strong the overall subjective opinions of the model builder are: if set equal to ∞, the opinions have no influence and the estimates are equal to the OLS ones; if λ is set equal to 0, the *a posteriori* estimates are equal to the *a priori* ones and the data have no influence on the estimates.

On the other side, again following the procedure by Banbura, Giannone and Reichlin (2010), the matrix $\mathbf{X}$ is extended by means of the following:

$$\underset{46\times 31}{\mathbf{X}_d} = \begin{bmatrix} \mathbf{J}_2 \otimes diag\langle \sigma_1,\ldots,\sigma_{15}\rangle / \lambda & \underset{15\cdot 2\times 1}{\mathbf{0}} \\ \underset{15\times 15\cdot 2}{\mathbf{0}} & \underset{15\times 1}{\mathbf{0}} \\ \underset{1\times 15\cdot 2}{\mathbf{0}} & kk \end{bmatrix} \quad (4)$$

where the diagonal[16] matrix $\mathbf{J}_2 = diag\langle 1,2\rangle$, and kk is a very large number (for instance 10000), indicating the complete uncertainty of the model builder about the values of intercepts.

To interpret the $\mathbf{Z}_d$ and $\mathbf{X}_d$ matrices, it can be noted that their first rows (the first two in $\mathbf{Z}_d$ and the first one in $\mathbf{X}_d$) impose the subjective opinions on the autoregression parameters; the following row imposes the opinions on the dispersion matrix of disturbances; and the last row represents the *a priori* opinion (non- informative) on the intercept. As examples, it can easily be shown that:

[15] In this case both Litterman (1986) and Banbura, Giannone and Reichlin (2010) set $\delta^{(1)} = 0$.

[16] The diagonal elements of the matrix $\mathbf{J}_2$ represent the degree of uncertainty of the model builder, which grows gradually as the lags increase.

$$E\left(\tilde{\varphi}_{ii}^{(j)}\right) = \delta_i^{(j)} \qquad \forall i, \forall j$$
$$E\left(\tilde{\varphi}_{ik}^{(j)}\right) = 0 \qquad \forall i, \forall k, \forall j, i \neq k \tag{5}$$
$$\ldots$$

Similarly, it can be shown how it is possible to obtain the diagonal elements (the others are taken equal to zero) of the *a priori* dispersion matrix of parameters, of the Minnesota type, when the latter are set in vectorized form.

Adding the dummy observations, the regression model (2) becomes:

$$\underset{n_* \times 15}{\mathbf{Z}_*} = \underset{n_* \times k}{\mathbf{X}_*} \cdot \underset{k \times 15}{\mathbf{B}} + \underset{n_* \times 15}{\mathbf{U}_*} \tag{6}$$

where $n_* = n + 46$, $\mathbf{Z}_* = \left[\mathbf{Z}' \, \mathbf{Z}_d'\right]'$, $\mathbf{X}_* = \left[\mathbf{X}' \, \mathbf{X}_d'\right]'$, and $\mathbf{U}_* = \left[\mathbf{U}' \, \mathbf{U}_d'\right]'$, with $\mathbf{U}_d$ matrix of the model disturbances, conform to $\mathbf{X}_d$.

With this addition the mean values of the autoregressive coefficients are shown to be equal to the OLS estimates of (6) and at the same time to coincide with the mean values of the Minnesota *a priori* distribution, as indicated by relations (5). Similarly for the parameters of the dispersion matrix. Note the importance of the result: the VAR model can be easily estimated because the order of the matrix to be inverted is computationally low, despite the high number of variables (even of the order of several dozens of units). At the same time, it is possible to take account of the subjective opinions of the model builder, as in the Minnesota distribution.

4. The a priori subjective opinions

Let's define now the structure of the a priori distribution for the VAR model built in Section 2. First of all, as used in the a priori distributions à la Litterman, the impacts of variables lagged twice are supposed to be zero. Thus: $E(\tilde{\varphi}_{ki}^{(2)}) = \delta_{ik}^{(2)} = 0, \quad \forall i, \forall k$.

As for the lag of one, all variables are supposed to possess a high persistence (which in many cases corresponds to the trend); so we set: $E(\tilde{\varphi}_{ii}^{(1)}) = \delta_{ii}^{(1)} = 1, \quad \forall i$. But, unlike what is usually done in the distributions of Minnesota type, we suppose that a priori also some variables may impact on others, as it happens in reality. In particular, it may be assumed:

- that the quantity of money influences exchange rate and prices in a basic way; therefore:
  - for Japan: $E(\tilde{\varphi}_{5,1}^{(1)}) = \delta_{1,5}^{(1)} = E(\tilde{\varphi}_{8,1}^{(1)}) = \delta_{1,8}^{(1)} = 0.3$
  - for the USA: $E(\tilde{\varphi}_{9,2}^{(1)}) = \delta_{2,9}^{(1)} = 0.3$
  - for the Eurozone: $E(\tilde{\varphi}_{6,3}^{(1)}) = \delta_{3,6}^{(1)} = E(\tilde{\varphi}_{10,3}^{(1)}) = \delta_{3,10}^{(1)} = 0.3$
  - for Cina: $E(\tilde{\varphi}_{7,4}^{(1)}) = \delta_{4,7}^{(1)} = E(\tilde{\varphi}_{11,4}^{(1)}) = \delta_{4,11}^{(1)} = 0.3$
- that the quantity of money influences GDP in a less marked way in all countries or monetary zones:
  - $E(\tilde{\varphi}_{12,1}^{(1)}) = \delta_{1,12}^{(1)} = E(\tilde{\varphi}_{13,2}^{(1)}) = \delta_{2,13}^{(1)} = = E(\tilde{\varphi}_{14,3}^{(1)}) = \delta_{3,14}^{(1)} = E(\tilde{\varphi}_{15,4}^{(1)}) = \delta_{4,15}^{(1)} = 0.1$
- that the exchange rate influences prices and GDP in all Countries or monetary zones:
  - $E(\tilde{\varphi}_{8,5}^{(1)}) = \delta_{5,8}^{(1)} = E(\tilde{\varphi}_{12,5}^{(1)}) = \delta_{5,12}^{(1)} = E(\tilde{\varphi}_{6,10}^{(1)}) = = \delta_{6,10}^{(1)} = E(\tilde{\varphi}_{14,6}^{(1)}) = \delta_{6,14}^{(1)} = 0.2$
  - $E(\tilde{\varphi}_{11,7}^{(1)}) = \delta_{7,11}^{(1)} = E(\tilde{\varphi}_{15,7}^{(1)}) = \delta_{7,15}^{(1)} = 0.2$
- that the remaining impacts are zero.

A low value17 is given for λ, 0.08, showing the strong degree of belief by which the opinions of the model builder are expressed. By constructing a univariate autoregressive model of fourth order for each variable, the following values for the standard deviations of disturbances are obtained:

$$\sigma_1 = 0.013, \sigma_2 = 0.045, \sigma_3 = 0.004, \sigma_4 = 0.010, \sigma_5 = 0.012,$$
$$\sigma_6 = 0.005, \sigma_7 = 0.005, \sigma_8 = 0.011, \sigma_9 = 0.039, \sigma_{10} = 0.023,$$
$$\sigma_{11} = 0.004, \sigma_{12} = 0.023, \sigma_{13} = 0.041, \sigma_{14} = 0.007, \sigma_{15} = 0.005$$

[17] Usually set to 0.10 by the users of Minnesota distribution.

5. Monetary shocks in the USA, the Eurozone, Japan and China

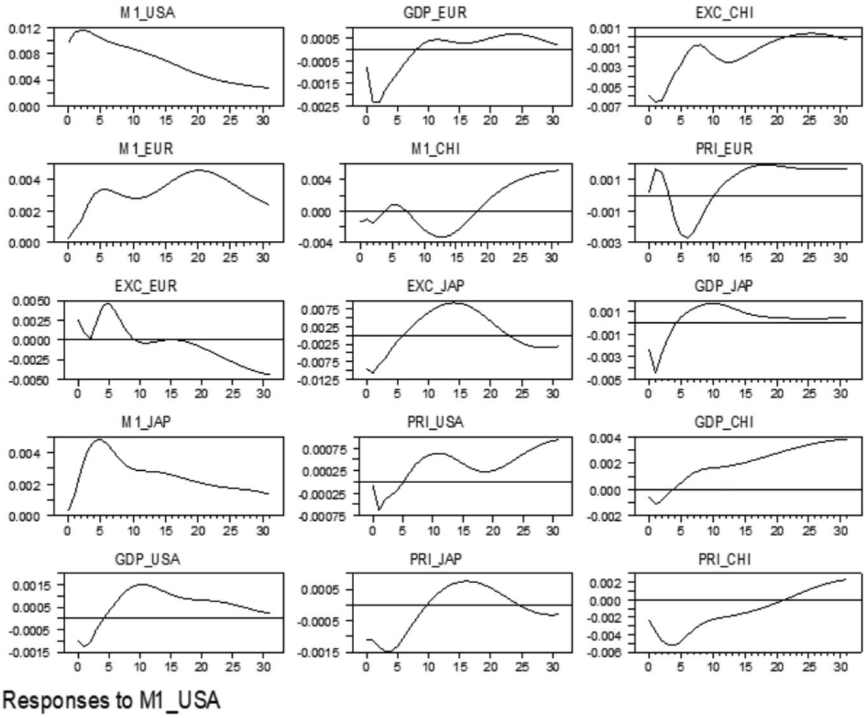
Although the VAR models are the reduced form of simultaneous structural models, the statistical significance of their parameters or the correlation between these is not of great interest. The usefulness of the VAR models concerns instead the overall effects on specific variables produced by shocks imposed on (the disturbances of) some of them. These effects are shown for a pretty long period of time and constitute the so-called impulse (the shock imposed on a variable) response (the effects on another) function (IRF).

This Section examines some IRFs that are considered interesting for the objectives of the paper. Figures 1, 2, 3 and 4 show the responses of the 15 variables in the system, over 32 quarters, to impulses applied to the quantity of money (M1) in the USA, the Eurozone, Japan and China, respectively. Shocks consist of one standard deviation (of disturbances) and in the figures the effects on variables are measured in terms of their own standard deviations.

5.1. Effects of a monetary shock in the USA

Figure 1 shows the IRF induced by a shock to US money. The GDP of this monetary zone (GDP-USA) initially falls and then grows steadily to its maximum (after 10 quarters). Then it returns to neutrality. The same trend is found in other analyses, one of which, among the main ones, is by Leeper, Sims, Zha, Hall and Bernanke (1996). As usual, there is the price-puzzle, consisting of price (PRI-USA) that falls simultaneously with the shock to money, and then grows in both the short and long term.

Figure 1. Impulse responses to a shock to US money



The sudden growth of US money is immediately reflected in a fall of the yen/dollar exchange rate (EXC-JAP). Japan's GDP declines significantly with the delay of one quarter but monetary authorities appear to intervene by increasing (albeit slowly) money and improving both the exchange rate and output. Moreover, Eurozone's monetary authorities appear to react much better than in Japan, as the euro/dollar exchange rate (EXC-EUR) is not particularly affected. Eurozone's output (GDP-EUR), on the other hand, is affected, as in Japan (GDP-JAP), with the minimum after two quarters. Eurozone's (PRI-EUR) and Japan's (PRI-JAP) prices possess similar trends.

The latter grows as the result of the increase in money, after having shown the price-puzzle: Eurozone's price falls to its minimum after six quarters since the monetary shock, while Japan's price reaches the minimum after three. Similarly, China's price (PRI-CHI) falls in the short term (the minimum is after three months), while its GDP increases slowly but steadily, in both the medium and long term.

5.2. Effects of a monetary shock in the Eurozone

Figure 2 shows the IRFs induced by a shock to Eurozone's money. It produces the expected effects on its own economy: after the initial increase, its M1 steadily declines until the simulation end; its GDP grows steadily up to its maximum over five quarters, and then returns to neutrality, which reaches after ten quarters. Likewise for price, which grows up to the maximum over seven quarters and returns to the normal path more slowly. So, the behaviour of Eurozone's economy after a shock to its money seems to be quite similar to that of US economy as the result of a monetary shock.

The effects on the other countries, which are very weak indeed, do not look like the expected ones. The euro in the medium term appreciates lightly against the dollar (EXC-EUR) and consequently in the period 4-8 quarters both output and price grow a little in the USA. In Japan, GDP only has a light positive initial jump and its price a negative one.

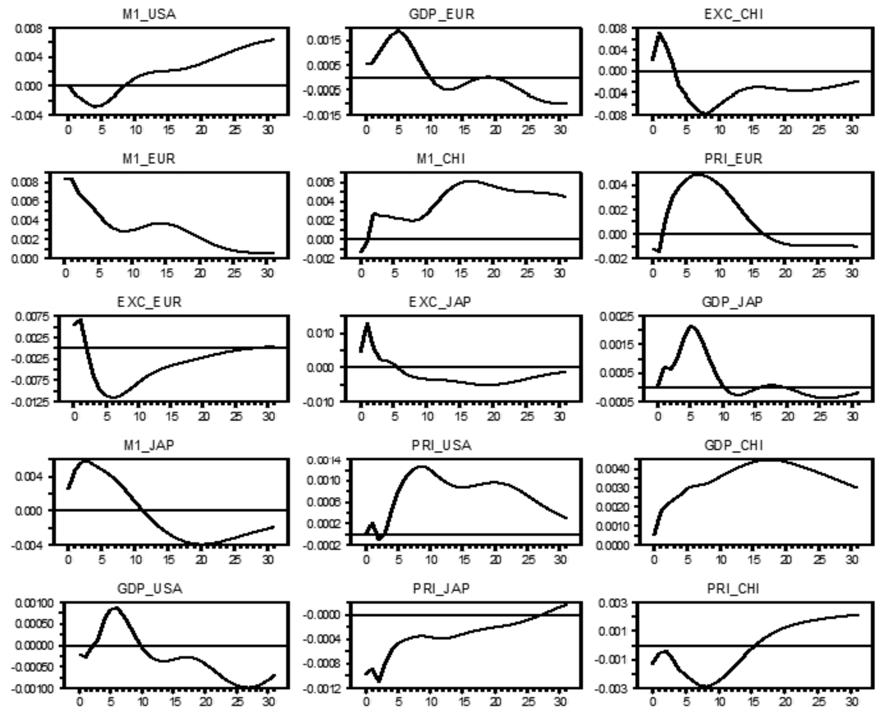
In all probability, there is a growth in consumption in the Eurozone and hence an increase in imports from the USA and Japan. Also China's GDP does not contract, but expands robustly and consistently in both the short and long term.

5.3. Effects of a monetary shock in Japan

A monetary shock in Japan (Figure 3) produces poor results on its own economy: its GDP initially declines and then grows, but lightly (3-7 quarters). Its price has a positive outbreak lasting one quarter only, and afterwards it returns to the normal path. Also the effects on US, Eurozone's and China's outputs are not substantial: initially zero, they become positive only in the medium term and over a short time. They reach the maximum after 10 quarters in the USA and after 13 in the Eurozone. China's GDP increases steadily, but slowly, up to the maximum after twenty-five quarters.

Prices in the USA and Eurozone behave accordingly: they start rising almost immediately in the Eurozone and after five quarters in the USA. The former reaches the maximum after four quarters while the latter reaches it slowly, after fourteen.

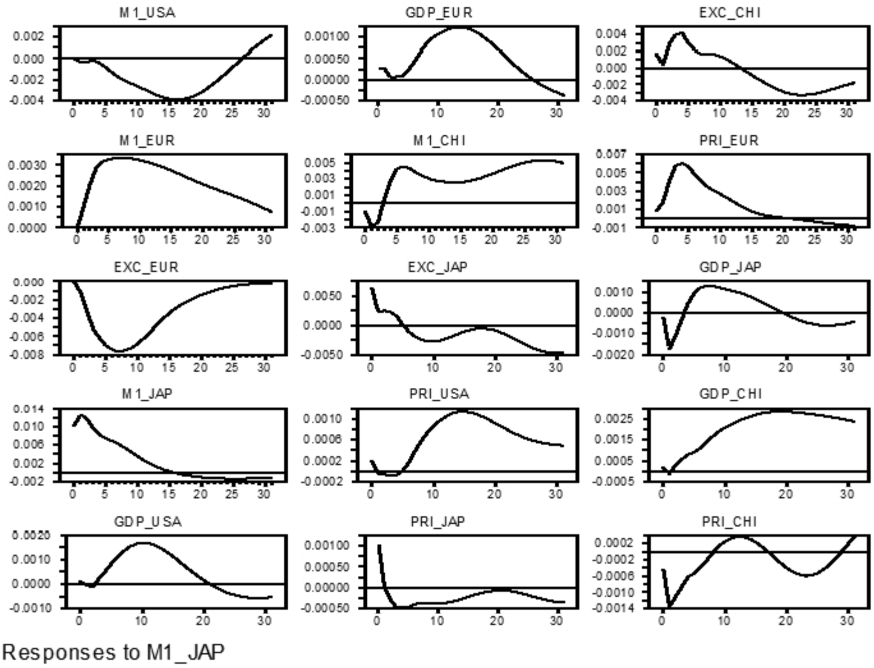
Figure 2. Impulse responses to a shock to money in the Eurozone



It is interesting to note that the monetary shock in Japan immediately lowers its exchange rate (EXC-JAP), which, however, returns to its normal path (and stays there) after four quarters. The effects on Eurozone's exchange rate are much slower: even if the euro begins to appreciate after one quarter, it reaches the maximum appreciation after seven. It returns to its normal path only in the long run, after thirty quarters.

To sum up, a standard monetary shock in Japan produces poor results. Probably the country would require very strong shocks to really improve its economic activity.

Figure 3. Impulse responses to a shock to money in Japan

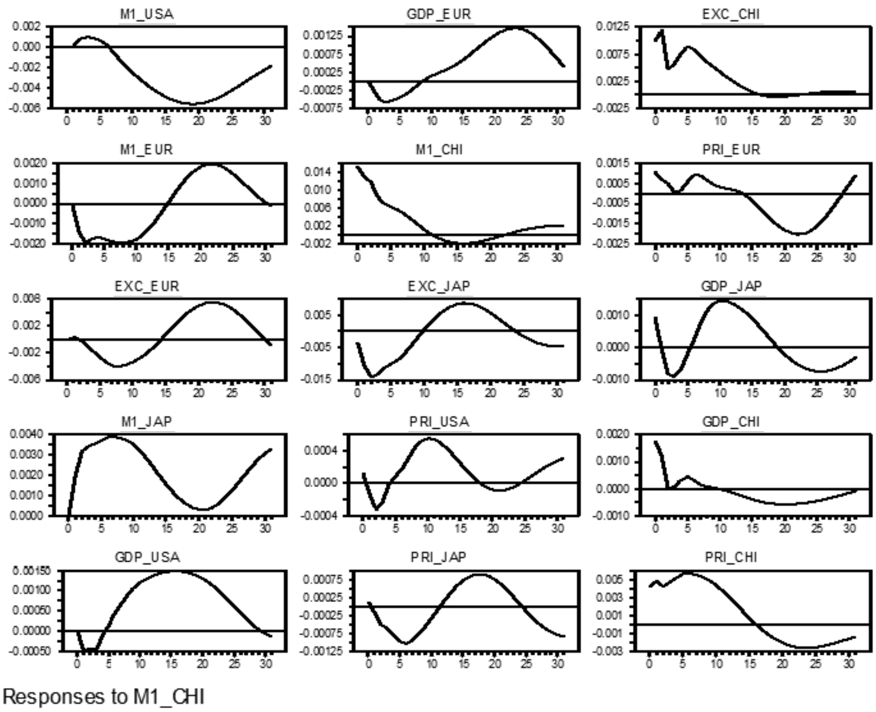


5.4. Effects of a monetary shock in China

The effects of a monetary shock in China (Figure 4) are relatively non-orthodox. First, it has a very marginal impact on its own GDP (GDP-CHI): it increases only in the first quarter after the shock, and then returns to its normal path. It is more effective on domestic price (PRI-CHI), which grows immediately and remains higher than normal over 15 quarters.

After a slight decline, evident in Japan but very low in the USA and the Eurozone, GDPs of the other three countries (or monetary zones) increase very lightly: quickly in Japan, for which output reaches the maximum after 11 quarters; slowly in the USA, with the maximum after 15 quarters; even more slowly in the Eurozone with the maximum after 23 quarters. Also price moves, even much lightly: it decreases in the short term in the USA (minimum after two quarters) and in Japan (minimum after six). Then it reaches the maximum, which is in the USA in the tenth quarter, and in Japan in the eighteenth. On the contrary, in the Eurozone price is not undergoing major alterations in the short-medium term.

Figure 4. Impulse responses to a shock to money in China



To sum up, the effects of a monetary shock in China seem to be very poor, for the country itself and other three countries.

6. The decomposition of the forecast-error variance

A second type of analysis that can be done through the VAR models concerns the decomposition, for each variable, of the forecast-error variance. This analysis makes it possible to determine the importance of each variable in explaining the variability of one of them. And since the forecast involves a long period of time, such an analysis is dynamic, in the sense that it is possible to see how this importance evolves over time, diminishing or widening. Figures 5 and 6 show the decompositions of the forecast-error variance made over 32 quarters ahead for some relevant variables: US, Eurozone's and Japan's outputs, as well as US and Eurozone's prices. They indicate how much of their variance, as a percentage, is due to US, Eurozone's and Japan's money and GDP.

To interpret these results well, it should be noted that in Figure 5 the percentages of explained variance are all between 0% and 10%, while in Figure 6 in only four cases they vary between 0% and 100%, and in two cases between 0% and 30%.

Let's first look at Figure 5. US output is highly influenced by US money, as expected, but also by Japan's money in the medium term, and not by that of the Eurozone.

This fact seems to confirm the implications of the impulse responses: US output responds positively to a Japanese monetary shock in the medium-long term (10-20 quarters) but poorly to a monetary shock in the Eurozone. The money of the latter, as expected, partly explains the variability (in forecast) of its own output, while does not affect the GDPs of the USA and Japan, if not lightly in the medium-long term.

Lags are those theoretically expected: US money has the greatest impact on US GDP after 12 quarters, and afterwards the impact remains constant, while Japanese money has a maximum that is more delayed.

Taking account of dimensions, Japanese money seems to affect more significantly the US output than the Japanese one, which appears, probably by virtue of monetary channels (through bonds, for instance), much less responsive. An impact of Japanese money that is similar, but dimensionally more limited, appears on Eurozone's GDP, as shown by the comparison of the first two panels high on the right in Figure 5.

US price is heavily influenced by monetary variations in the USA themselves, while it is poorly affected by what is occurring outside, in the Eurozone and Japan. A similar fact happens to Eurozone's price, which however seems not to be affected by changes of its own money.

Furthermore, it does not seem that the Japanese money variations, even if impacting on outputs of the USA and Eurozone, have any effect on their prices.

Figure 5. Percentages of the forecast-error variance due to the US, Eurozone's and Japan's money for five variables

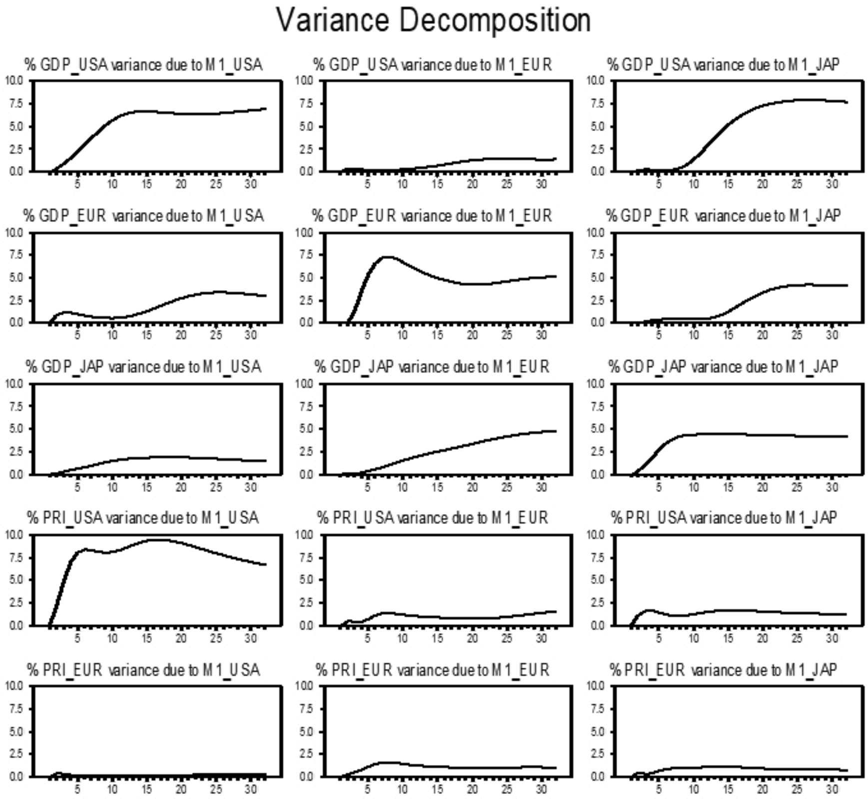
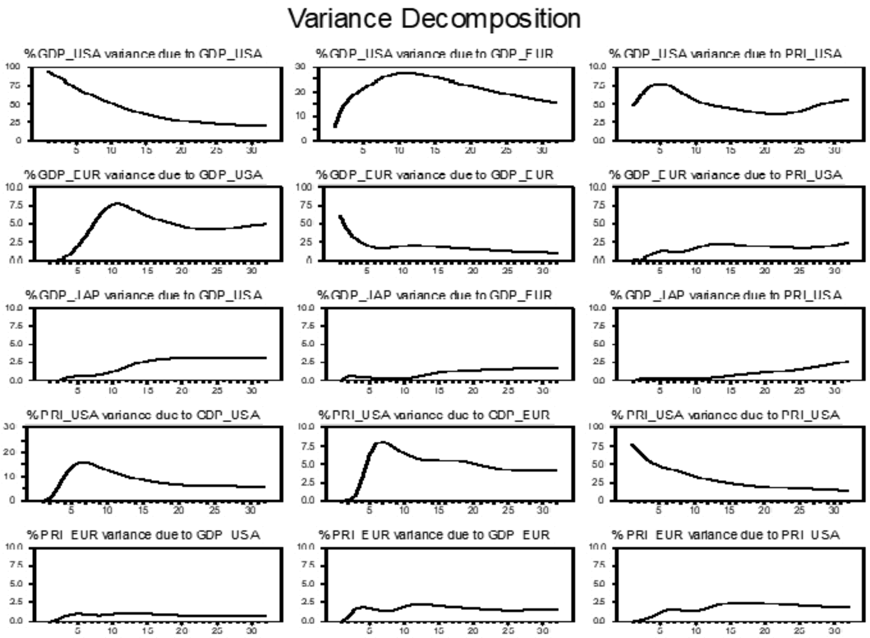


Figure 6 analyses the variance percentages that are attributable no longer to money, but to US and Eurozone's GDPs as well as US price. This is related to US GDP with a light impact, as expected, whilst it appears to be largely irrelevant to the GDPs of the Eurozone and Japan. Instead, US and Eurozone's GDPs lightly affect each other, with a maximum of about ten quarters in both directions. Strangely, Eurozone's price is not affected by US GDP, while US price is lightly affected by Eurozone's output in the short term (5-7 quarters). This latter influence goes on in the medium-long term.

To sum up, the figure seems to show that prices interact very weakly between the USA, the Eurozone and Japan. A similar weakness is perceived in the links of their outputs, with the possible exception of a light impact between the output of the USA and the Eurozone.

Figure 6. Percentages of the forecast-error variance due to the US and Eurozone's GDP, as well as to US price, for five variables



China is not considered in the decomposition of the forecast-error variance as it has proven not to be influenced economically by the other countries, nor influencing them.

7. Concluding remarks

The effects of the international monetary shocks can be easily analysed by means of the VAR models, which do not require the specification of particular relationships between variables. In the paper, a VAR model containing four variables for four countries is built. The variables are: money, national-currency/US-dollar exchange rate, consumer price index, and GDP; the countries (or monetary zones): the USA, the Eurozone, Japan and China. Using the Cholesky factorization, the model, despite the high number of variables, can be estimated with the ordinary least squares.

But estimation has to be of the Bayesian type in order to be carried out by exploiting the subjective opinions of the model builder. This can be done by adding some dummy observations to the real ones. The procedure for this addition is described in Section 4 and its implementation to the case of this paper in Section 5. In the previous literature variables are supposed a priori not to affect the others, but only themselves at lag one. In this paper, on the contrary, some corrections have been made to the methodology to allow each variable to interact a priori with the others.

The analysis of the effects of monetary shocks and the decomposition of the forecast-error variance has produced sometimes expected, sometime unexpected results. A monetary shock in the USA has expected effects: an improvement in the US GDP and a worsening in the Eurozone's and Japan's outputs; increase of price in the USA and decrease in the other countries. The only exception seems to be China's GDP, which grows slowly but steadily after the shock. On the contrary, a monetary shock in the Eurozone, while increasing the output of its member countries, does not diminish that of the USA, nor those of Japan and China. Even US price is somehow affected by a light increase. Instead, Japan's and China's prices show a first negative impact, which is absorbed only in the medium (for China) or long term (for Japan).

The effects of a monetary shock in Japan are poor for its economy. In the USA and Eurozone they are the opposite of those expected: outputs improve and accordingly prices increase. China's GDP also increases, slowly but steadily. The monetary shock in Japan immediately lowers its exchange rate, which returns after a year to its normal path. In China, a monetary shock that occurs in the country does not produce significant effects on its GDP. It is more effective on its domestic price, which grows immediately and remains higher than normal over 15 quarters. The effects on economies of the other countries are very light.

US output is heavily influenced, in dynamic-regression terms, by US money, as expected; very lightly by GDP of Japan in the medium term; and not by that of the Eurozone. The money of the latter, as expected, partly explains the variability of output of the Eurozone itself, while not affecting the GDPs of the USA and Japan, if not lightly in the medium-long term.

In forecast terms, US and Eurozone's GDPs seem only lightly influence each other. US price is heavily affected by monetary variations in the USA themselves, while it is not affected by what is happening in the Eurozone and Japan. Strangely enough, Eurozone's price seems not to be influenced by changes of its own money, nor by US GDP, while US price is somewhat affected in the short term by Eurozone's output. This latter influence continues in the medium-long term.

References

Banbura M., Giannone D., Reichlin L. (2010), *Large Bayesian vector autoregressions*, Journal of Applied Econometrics, 25, 1, 71–92.

Christiano L.J., Eichenbaum M., Evans C.L. (1999), *Monetary policy shocks: What have we learned and to what end?*, Handbook of Macroeconomics, ed. by Taylor J. B., Woodford M., vol. 1, chap. 2, 65–148, Elsevier.

Eggertsson G.B., Woodford M. (2003), *The zero bound on interest rates and optimal monetary policy,* Brookings Papers on Economic Activity, 2003, 1, 139-211.

Eichenbaum M., Evans C.L. (1995), *Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates*, The Quarterly Journal of Economics, 110, 4, 975–1009.

Eichengreen B., Sachs J. (1985), *Exchange Rates and Economic Recovery in the 1930s*, Journal of Economic History, 45, 925-946.

Eichengreen B., Sachs J. (1986), *Competitive Devaluation and the Great Depression*, Economics Letters, 22, 67-71.

Korineck A. (2013), *Capital Account Intervention and Currency Wars*, Centre for Economic Policy Research.

Leeper E.M., Sims C.A., Zha T., Hall R.E., Bernanke B.S. (1996), *What Does Monetary Policy Do?*, Brookings Papers on Economic Activity, 2, 1-78.

Litterman R. (1986), *Forecasting With Bayesian Vector Autoregressions – Five Years of Experience*, Journal of Business and Economic Statistics, 4, 25–38.

OECD (2016), Economic Outlook: Statistics and Projections, Paris, OECD Publishing

Rodrik D. (2008), *The real exchange rate and economic growth: Theory and evidence*, Brookings Papers on Economic Activity, 4, 365-412.

Sims C.A. (1980), *Macroeconomics and Reality*, Econometrica, 48, 1-48.

G. & L. E. R., Vol. 22 No. 2, 2018

Giordano Bruno [*]

THE SYSTEMIC THINKING OF BRUNO DE FINETTI

Abstract

The *mechanistic* scientific approach, based on principles such us that for which the microscopic world is simpler than the macroscopic one, and that the macroscopic one can be expressed through the strict knowledge of infinite details, it has been challenged by many questions, one of which is represented by the so-called "The Three-Body Problem" (Barrow-Green, J., 1996).

New cognitive approaches have been introduced based, just to name a few, on the theoretical role of the observer, on non-linearity, on principles of uncertainty, on constructivism, on the systemic point of view and on emergence.

Here, I wish to emphasize how the treatment of uncertainty, developed according by Bruno de Finetti's vision, fits perfectly in these types of approaches, rather he anticipates them, stopping to highlight what is the essential role of one who makes probability assessments, which we will call the *observer*.

Unlike objectivist orientations, in particular those that refer to a purely statistical view of the reality, Bruno de Finetti had, in the last century, the fundamental intuition that the probability assessments of an event do not represent anything else if not *the degree of believe* who a *coherent individual*

[*] ISIA Roma Design, piazza della Maddalena, n. 53, Rome, Italy, phone: +39-06-6796195. *E-mail address*: gibrun84@gmail.com

has in realizing that event. This degree of believe is determined on the basis of the *information* that the individual possesses regarding that event and in this sense the individual plays the role of observer, as happens in the modern *theory of complexity*.

de Finetti taught us that the logic of uncertainty explores the context of the possible, accepting the condition of being unsuitable for making *predictions*, on the contrary providing the tools to assign probabilities to events, so as to be able to make *previsions* and take decisions.

Traditionally, the concept of probability has been considered as a consequence of our *ignorance*, of our limitations. In the same way, the uncertainty's principles were considered by mechanistic thinking as a limitation of our knowledge of infinite details.

Now we can instead consider the probability as our *tool* to describe the nature.

The methodology followed by de Finetti is based on the *coherence* in the assignment of probabilities to single event and families of single events.

The coherence is the essential tool that allows us to make probability assessments and, in particular, to update the assessment of an event when a new (or supposed) information becomes usable (of interest).

.

JEL CLASSIFICATION: C00.
KEYWORDS: Uncertainty, Nonlinearity, Systemics, Emergence, Complexity, Information, Prevision, Coherence, Observer.

We desire *truth*, and find within ourselves only *uncertainty*

Blaise Pascal, *Thoughts.*
The Harvard Classics. 1909–14.

It would therefore seem natural that the customary modes of thinking, reasoning and deciding should hinge explicitly and systematically on the factor *uncertainty* as the conceptually pre-eminent and determinative element.

Bruno de Finetti, *Theory of Probability: A Critical Introductory Treatment*, 2 volumes (Translated by A. Machi and A. Smith). Wiley, London (1974).

1. Coherent probabilities assignments and emergence

I propose to show how the problem of a *coherent* assignment of probabilities to a family of event, by an *observer*, is related to *emergence*, in the way it is considered in the modern science.

I remember that *emergence* (Corning 2002, Minati and Pessa 2002, Pessa 1998, 2002), in this context, may be considered as a process of formation of new *collective entities* (which require a level description different from that used for the single elements), as swarms, flocks, traffic, industrial districts, markets; *collective effects*, such as super conductivity, iron-magnetism, laser effect. These entities are determined by *coherent* behavior (as perceived by an observer) of interacting components.

Referring to the cognitive model of (Anderson 1983) used by the observer, emergence may be seen as a process depending only by the observer, considering that collective properties *emerge* at a higher level (ie more abstract) than to the one he used to treat the components. Collective properties are detected by the observer as new, depending on the cognitive model assumed, suitable for establishing the presence of coherence.

But more, the subjective conception of de Finetti's probability is all built as a *logically open cognitive system*. In fact, the observer (the evaluator) plays an *active* role, sensitive to the *context*, who moves in a *non-objectivistic context*, is *process-oriented*, *learns dynamically* from experience in an *inductive* way, is an *integrating part of the system* and *generates its existence*, finally he can *choose the rules* to use (Minati 2004).

I will start from this last point to illustrate what has been said. In assigning a measure of uncertainty to a family of events, of which we do not know the truth value, the observer may follow two paths, depending on the context in which he finds himself (or to which he chooses to refer).

The first consists in introducing a *natural order* among the events: if we suppose, as it is legitimate, that an event compared with another may be or *more possible* or *less possible* or *equally possible* than this, he can easily construct a *relation of order* of *not less possible* within the family of events considered.

This relation has the following properties:

An uncertain event E_h of the family, is always more possible than the impossible event and less possible than the certain event and is no less possible than itself.

If an event E_h is no less possible than $E_{k,}$ then E_k cannot be no less possible than $E_{h,}$ unless E_h and E_k are equally possible.

If an event E_h is no less possible than E_k and E_k of $E_{j,}$ then E_h is no less possible than E_j.

If an event E is incompatible with E_k and with E_j and E_k is no less possible than $E_{j,}$ then the union of E with E_k is no less possible than the union of E with E_j.

In this way, the observer introduces a *qualitative measure* of the uncertainty of an event.

It should be noted, in passing, that in the particular case of dealing with a certain event partition in cases judged to be equally possible, it follows that the introduced qualitative order immediately translates in a quantitative measure of their uncertainty as a ratio between the number of favorable cases and the number of possible cases.

If, then, a further property related to conditional events is considered:

The events E_h and E_k imply E, then E_h/E is not less possible than E_k/E if E_h is not less possible than E_k ;

Its introduction together with the previous, allows to *qualitatively* develop the whole theory of probability (de Finetti 1937).

The second way is that for which the observer explains his own *degree of believe* in occurrence of an event by means a quantitative assessment.

To do so, it needs a measuring *tool* that is suitable for the purpose. Bruno de Finetti proposed two equivalent measurement criteria: that of the *bet* and that of the *penalty* (de Finetti 1974).

I recall here only that relating to bet, which appears more natural since it reflects historically what happened in the construction of the probability theory.

Suppose we have to bet a certain amount to win another one when a determined event occurs. This is what common happens in all betting fields.

Then make a *bet* on an event E means that you are willing to pay a *part* of a certain sum *S* (>0), which we can indicate with *pS*, to receive *S* if E occurs and *0* in the opposite case.

If we introduce the G_E *gain* function with respect to that given bet, we will get the following schematic:

$$G_E = \begin{cases} S-pS, & \text{if E occurs} \\ -pS, & \text{if E does not occurr .} \end{cases}$$

It should be noted that it is quite clear that in a bet that each of us has the goal of maximizing our own gain, so this fact could lead us to distort our assessment; how do defend ourselves from this to prevent it bet tool becomes arbitrary and therefore ineffective for the purpose to which it must respond?

Meanwhile, it must be ensured that Those who bets *pS* to receive *S* (from the other bettor) if E occurs, must likewise be willing to pay *S* to receive *pS* if E occurs, ie to exchange the terms of the bet with the other bettor, this will ensure that the assessment made by the individual reflects his or he degree of believe without being influenced by the desire to make a major gain, which otherwise could be achieve by the other. For example, suppose that in a bet on E, I think you can pay *70* to receive *100* in case of E occurs, I might think about increasing my hypothetical gain (in this case, equal to *30*), declaring to be willing to pay *40*; but if I am willing to exchange bet odds with the other bettor then my hypothetical gain could significantly decrease (*-60*)!

But this is not enough, it must ensure that the possible values of the G_E gain are not both of the same sign, because only in that case would there be a certain win or loss, regardless of the occurrence or not of E. With the consequence that only in such a case an individual would accept to bet. This condition was appropriately called by de Finetti *coherence.*

The coherence in a bet on an event E as it is well known establish that, given *S* equal to *1*, (but it is also valid for $S \neq 1$) in any case the price *p* that an individual is willing to pay to receive *1* if E occurs is always between *0* and *1*. Moreover, the same coherence imposes that if E is *certain,* then *p* must necessarily be *1* and if E is *impossible,* then must necessarily be *0*.

It should be noted, however, immediately that $p=1$ not implies E *certain*, nor does $p=0$ implie E *impossible* (for further details see de Finetti 1974).

So, following de Finetti, the probability of an event E (or the the *numerical measure* of uncertainty on E) is the price *p* (real number) that an individual is willing to pay in a *coherent bet* to receive *1* if E occurs and *0* otherwise.

But just a bet will be coherent only if the *observer* (the individual) who evaluates the uncertainty of E make it so.

There will therefore exist, whith respect to E infinite coherent assessments, provided they are between 0 and 1!

How then will the observer choose one? He will have to rely on the *information* he has about E and express this information through a number.

Naturally, the more the information will be *rich,* the more the individual will be less doubtful in choosing one among infinite numbers!

Now in measuring information, an individual, as always happens, does not have an *objective* evaluation criterion: in fact, personal beliefs, feelings and all those characteristics that contribute to the formation of a judgement will intervene. So, he will only have to formulate, we could say so, objectively how much he assesses subjectively!

This is how de Finetti expresses in *L'invenzione della verità* (2006):

"Any behaviors…in sense of believing that the occurrence of this or that event does is plausible therefore depends only on a feeling, on that same feeling that must honestly present as the true starting point, and that someone prefers instead to ignore and modestly hide behind a barricade of logical devices as cumbersome as they are empty".

Obviously, in some cases, it will be much easier: for example, if an individual judges equally possible the occurrence of E and that of its negation E^c, to both will assign the value of probability 1/2 (but are we sure that judge equally possible two *real* events would be so simple and natural?). While, if He judges five times more probable E compared to E^c, then it will assign to E probability 5/6 and to E^c 1/6 .

In other cases it may resort to assessments based on the relative frequency, but only, as de Finetti has well specified, when the event E is part of a family of *exchangeable* events, that is when the assessment of probability of any n-tuple of family events considered depends only on the number of fixed events and not of particular fixed events; in short: it depends only on *how many* events and not *which* they are considered (de Finetti 1974)!

For example, in the extraction with return of balls from an urn of unknown composition, that is, for wich we know the total number of balls but not the percentage of *red* ones, an individual who want to evaluate the probability to get a *red* ball at n-th extraction, having been extracted (n-1)/3 red on (n-1), it could estimate equal to (n-1)/3 the probability that the red ball comes out at the n-th extraction, because the events of the family considered are exchangeable (we only care *how many* red balls are extracted). But if we would like to evaluate the probability of a certain boxer victory at the 101-th match of his career, knowing that he won 85 on previous 100, would that be enough to estimate this probability equal to 85/100? Obviously no, because it could, in the worst case, have lost all the last 15 matches, and the degree of believe in him… may not be so high!

1.1. The case of random numbers and complex phenomena

I have done so far considerations relating to single events or to families of analogous events. Often however, we run into random phenomena, which in some cases can be described by *random numbers*. Such as the X number of fatal car accidents in a year relating to people who did not wear a seat belt, for which one can consider events of the type (X = n). Or in more complex phenomena, with respect to which we may be interested in such events: the amount of rainfall in a given Italian city in the next year will be greater than that of the past year, or the average humidity will be lower, or the measured smog level will be higher, or fuel consumption will be equal. In the first case, that of random numbers, it is possible to formulate various probabilistic models that allow us to evaluate the probability of any event related to them, but we must not forget that these evaluations are not objective, as they apparently may seem, because given the model is enough apply formulas to get them; vice versa it is always the observer who chooses, based on his information, the model he considers most suitable to describe the phenomenon considered. What does it happen, instead, in the second case?

Again, the systemic aspects of the approach adopted by de Finetti give an exemplar answer.

I remember that in the classical approach (commonly used in applications) relating to a random phenomenon we proceed as follows: first we define a Ω space of the results or elementary cases possible, or we construct a certain event partition, then we assign a probability to each of these cases (or constituents) and since any event related to that phenomenon can be obtained as a union of constituents, a probability is attributed to each of them in a linear manner. I note that of course the problem of how to assign probabilities to constituents remains open!

On the contrary, de Finetti bases his conception on the fact that every event is single and for each of them we can express our degree of believe through a qualitative or quantitative evaluation. If we then find ourselves in front of a single event and have evaluated its probability, we need to assign probabilities to further events how should we behave?

Several cases may arise. If you are dealing with a family of events E_i that forms a partition of the certain event, then for coherence the sum of the probabilities of the single E_i must be equal to 1 and the probability of the union of n incompatible events must be equal to the sum of the single probabilities.

If, again, there are n events and a coherent probability assignment, then the probability of an event that depends linearly on the first n is determined.

n the cases, instead, in which between the events E_{ii} considered and a new event E there is a logical connection, then always for the coherence it must proceed in this way: we build the constituents relatively to the given family (that is all the intersections possible between events so that in each appears one of them or its contrary, e.g. $E_1 \cap E_2 \cap \ldots \cap E_h \cap E^c_{h+1} \cap \ldots \cap E^c_{n-1} \cap E^c_n$), then two events are identified E 'and E ", which are respectively the maximum event of all the constituents implying E and the minimum event of all the constituents implied by E: necessarily the probability of E must be within the closed interval [P (E '), P (E ")].

Even in this situation, therefore, the probability of E is not univocally determined, it can simply shrink the interval [0,1] within which the observer can evaluate it to be coherent. then having indicated with E' and E", two events which are respectively the maximum event union of all the constituents which implies the event E, and the minimum event union of all the constituents which are implicated by the event E: the probability of E may be included necessarily in the closed range [P(E'), P(E")]. Even in this situation, therefore, the probability of E is not univocally determined, it can simply shrink the interval [0,1] within which the observer can evaluate it to be coherent.

But the most interesting situation occurs when dealing with individual events that are examined starting with the first and then the others afterwards. To each of these the observer, based on his own information, assigns a coherent probability value, i.e. between 0 and 1. In this case, however, there may be logical links between events that have not been considered or, otherwise, referred to information was received only when all the events were introduced, then how to check whether the assessment is overall coherent?

It will be necessary to construct the constituents starting from the events considered, and since each of these latter will result in the union of some of the constituents obtained, its probability must be equal to the sum of the probabilities (not yet determined, therefore let's say so unknown) of these constituents. In this way we will obtain a system of n equations in s ($\leq 2n$) unknown x_i, with the constraints $x_1 + x_2 + \ldots + x_s = 1$ and $x_i \geq 0$. If there is a complete solution of the system then the evaluation given may be said to be coherent.

It is interesting to note that the system may not allow solutions, in which case the evaluation would be incoherent; that can admit a single solution, but that there can also be more solutions, that is a set of different evaluations all coherent. If exists an s-tuple solution of the system, the given assessment could be defined as coherent.

I will clarify this last interesting aspect through examples.

Given three events A, B, C and one observer would have evaluated their probabilities as follows P(A)= 1/2, P(B) = 2/5, P(C) = 1/5 (obviously, each of them represents a coherent assessment!).

Let it be known then that A B C = Φ (with A B C = $A \cap B \cap C$).

So, the possible constituents are:

$Q_1 = A^C B C$, $Q_2 = A B^C C$, $Q_3 = A B C^C$, $Q_4 = A B^C C^C$, $Q_5 = A^C B C^C$,
$Q_6 = A^C B^C C$, $Q_7 = A^C B^C C^C$.

To establish then whether, under the given conditions, the overall evaluation expressed by P(A) = 1/2, P(B) = 2/5, P(C) = 1/5 would result coherent, it is necessary to establish whether the following system admits at least one solution:

$$\begin{cases} x_2 + x_3 + x_4 = 1/2 \\ x_1 + x_3 + x_5 = 2/5 \\ x_1 + x_2 + x_6 = 1/5 \end{cases}$$

with $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1$ and $x_i \geq 0$, $i = 1, 2, \ldots, 7$.

If it is posed, as it is legitimate to do, $x_1 = 0$, $x_3 = 0$, $x_6 = 0$, with simple steps we get the following system solution

$x_1 = 0$, $x_2 = 1/5$, $x_3 = 0$, $x_4 = 3/10$, $x_5 = 2/5$, $x_6 = 0$, $x_7 = 1/10$;

therefore, the allocation of assigned probabilities to the events A, B, C determine a coherent overall assessment!

Note that if we put $x_2 = 0$, $x_3 = 0$, $x_6 = 0$, we would have obtained a different one system solution

$x_1 = 1/5$, $x_2 = 0$, $x_3 = 0$, $x_4 = 1/2$, $x_5 = 1/5$, $x_6 = 0$, $x_7 = 1/10$;

and also, in this case, the overall assessment would have been coherent!

Not only, but if initially the observer had evaluated $P(A)=\alpha$, $P(B)=\beta$, $P(C)=\gamma$, with the condition A B C = Φ, we would have obtained the following system

$$\begin{cases} x_2 + x_3 + x_4 = \alpha \\ x_1 + x_3 + x_5 = \beta \\ x_1 + x_2 + x_6 = \gamma \end{cases}$$

with $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 1$; $x_i \geq 0$, $i = 1, 2, \ldots, 7$,
and $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$.

With the choices of $x_i = 0$, made in the two cases examined above, we would have had respectively the following solution:

$x_1 = 0, x_2 = \gamma, x_3 = 0, x_4 = \alpha-\gamma, x_5 = \beta, x_6 = 0, x_7 = 1-(\alpha+\beta)$,

which would give rise to a coherent overall assessment, as long as would be $\alpha \geq \gamma$ and $\alpha+\beta \leq 1$;

$x_1 = \gamma, x_2 = 0, x_3 = 0, x_4 = \alpha, x_5 = \beta - \gamma, x_6 = 0, x_7 = 1-(\alpha+\beta)$,

that would give place to a total coherent assessment, provided it is $\beta \geq \gamma$ and $\alpha+\beta \leq 1$.

2. Learning by experience

I now turn to examine the additional aspects of the initial assumption, in particular for as regards learning inductively by experience, referring to conditional events.

I recall that, given any two events E and H ($H \neq \Phi$), we can consider a new entity E/H, to which the name of conditional event is assigned, which has the following meaning:

	TRUE,	if H true and E true
E/H =	FALSE,	if H true and E false
	UNDETERMINED,	if H false

So, if we want to make a (conditional) bet relative to E/H we will, according to de Finetti[1], behave as follow:

		1	if H true and E true
PAY p	TO RECEIVE	0	if H true and E false
		p	if H false

From this definition of conditional bet it follows that we can measure the uncertainty of the E/H event through the price p that a coherent individual is willing to pay with the previous outcomes and this measure is called conditional probability and denoted P (E/H).

de Finetti proves, moreover, how the conditional probability, thus introduced, verifies all the properties of a probability and, finally, as the only possible reading of P (E H) is that of the probability of E supposed to be true H.

[1]Bruno de Finetti, *Theory of Probability: A Critical Introductory Treatment*, 2 volumes (Translated by A. Machi and A. Smith). Wiley, London (1974).

In particular, he shows how the simple (and natural) condition of coherence: the unevenly negative random gain in a set of bets, leads to the *multiplication theorem of probability*:

given two events, any E and H ($H \neq \Phi$) we have

$P(E \cap H) = P(H)\, P(E/H)$;

and its corollary, the *Bayes theorem* (also with $E \neq \Phi$):

$P(H/E) = K\, P(H)\, P(E/H)$, with $K = 1/P(E)$, $P(E) \neq 0$,

which tells us that the conditional probability of H given E is equal to the product of the unconditional probability of H for the conditional probability of E given H, within a proportionality's not null K factor.

Let us now focus on the meaning of Bayes' theorem. In reality, it tells us much more if we give a reading of this kind: let it be an event that represents the outcome of an experiment of a given random phenomenon and let H be a hypothesis concerning the same phenomenon, then the theorem states that the conditional probability of the hypothesis H given E is proportional to the unconditional probability of the hypothesis H multiplied for the conditional probability of E given H.

To make it clearer what we have expressed, we use the classic example of the urn of *unknown composition*. In other words, let us consider an urn of which we know that it contains N balls, but of which the percentage of red balls is not known (being able, therefore, to be present in the urn from 0 to N red balls). The event E represents a possible result related to the extraction of n balls from the urn (for example, without restitution): h red balls on n, and event H is the hypothesis: r balls are present in the urn red on N. Then the Bayes' theorem let us to evaluate the conditional probability of the hypothesis H given the event E (we may call it *final probability*), through the unconditional probability of H (we may call it *initial probability*) and the conditional probability of E/H, such as of E supposed as true H (it may be calculated easily and we may call *likelihood*), always less than the proportionality factor K.

Ultimately, Bayes theorem tells us how we need to update our probability assessment in the presence of new information (or rather, suppose that new information becomes known to us):

final probability = K x *initial probability* x *likelihood.*

Observe that the initial and final have, in this context, the only meaning respectively of before and after it becomes known (suppose we know) E. Naturally, in the same way the final probability of hypothesis H^c can be evaluated, resulting

$P(H^c/E) = K\, P(H^c)\, P(E/H^c)$, with $K = 1/P(E)$, $P(E) \neq 0$.

And, more generally, if we want to evaluate the final probability of m different hypotheses Hj, forming a certain event partition, we have the following expression of Bayes theorem:

$P(H_j/E) = K\, P(H_j)\, P(E/H_j)$, $j=1,2,\dots,m$;

with $K = 1/P(E)$, $P(E) = \sum_{j=1,2,\dots,m} P(H_j)\, P(E/H_j)$.

Returning to our example, the Hj would represent the possible hypotheses on the composition of the urn and we, after having given an initial evaluation, we will be able to express a final through the Bayes theorem.

In general, in the objectivist settings of probability (the classical one and the frequentist one), since the probability evaluations are essentially numerical relationships, the observer (the individual who evaluates) has the only task of performing calculations well, based on situations of symmetries or repetitions of a given phenomenon (apparently, however, because in any case precise and subjective choices have been made beforehand: the possible cases considered as equiprobable and the repeated tests considered equiprobable and independent!). In these conceptions the Bayes theorem loses its meaning and its intrinsic value and assumes the role of a purely mathematical result.

In the subjectivist conception of Bruno de Finetti, on the contrary, he best expresses the meaning of how one can (and must) learn from the experience!

It is the same logical process that leads a doctor, to take another example, to make a diagnosis. He, in assessing the presence or absence of a disease in a patient, first visits and expresses a first idea about the disease, then makes him perform a series of diagnostic-instrumental tests, the results of which he uses to establish his presence or not of the disease.

It should be noted, incidentally, that this procedure seems to be part of the logic of the certain, while vice versa more generally it concerns the ambit of the logic of the uncertain: always and only of probability evaluations it is, even if fortunately often the final probabilities of ascertaining the disease are close to 1 or 0 (just think of the fact that in addition to the personal evaluation, clinical trials can also be affected by an error)! In any case, when we go to a doctor, the event: I am suffering from a given disease, it is always possible, neither certain, nor impossible!

Often in applications to estimate the value of some parameter related to a given random phenomenon, the so-called *maximum likelihood method* is used: in short and schematizing it is a matter of evaluating the different values of the probabilities of E (or the probability density, in the case of a continuous parameter) subject to the hypotheses Hj (the l*ikelihoods*) and to obtain the maximum value, which is chosen as an *estimation*.

Returning, once again, to our example of the urn of unknown composition we calculate the $P(E/H_j)$, for each j, and then the largest one of all is found; if this, for example, is $P(E/H_3)$ we say that hypothesis H_3 is an estimate of the true composition of the urn. Exemplifying further, in an urn there are 10 balls, of which it is not known how many are red ones. A sampling is carried out, that is a repeated extraction (with restitution) for a total of 5 balls and of these if there are 3 red ones and both E this event. The $P(E/H_j)$, j = 0,1, ..., 10 are evaluated, obtaining

$P(E/H) = 5!/3!2!\ (j^3/10)\ (10\text{-}j)^2/10^2$;

and distinctly

$P(E/H_0) = 0$

$P(E/H_1) = 0.0081$

$P(E/H_2) = 0.0512$

$P(E/H_3) = 0.1323$

$P(E/H_4) = 0.2304$

$P(E/H_5) = 0.3125$

$P(E/H_6) = 0.3456$

$P(E/H_7) = 0.3087$

$P(E/H_8) = 0.2048$

$P(E/H_9) = 0.0729$

$P(E/H_{10}) = 0$.

So, since the maximum value is 0.3456, which is relative to the hypothesis H_6, it can be deduced that the estimate (of maximum likelihood) for the unknown composition of the urn is 6 red balls out of 10.

This methodology has two contraindications, both logical ones!

The first is to make an estimate relative to hypotheses, using the inverse conditional probabilities $P(E/H_j)$ with respect to those that should be more correctly compared: $P(E/H_j)$; only through these, in a logically unequivocal manner, is it possible to establish which of the hypotheses of composition of the urn is the most probable, observed (supposed to observe) E, and take it as an estimation of the composition!

The second is that in this procedure the probabilities of the hypotheses $P(H_j)$ do not attend, which can always be evaluated and whose non-consideration can be absolutely harmful for the conclusions reached.

In fact, also in the previous example, if we used (as it is necessary) the Bayes theorem, in the evaluation of $P(E/H_j)$ the $P(H_j)$ would enter, and any distribution of these different from the equiprobability, could make maximum the final probability of one of the H_j different from H_6. If, for some reason, it was known that it is much more likely that as many red balls of another color than the other hypotheses have been placed in the urn, that is in our case we had, for example $P(H_5) = 0.6$, $P(H_4)=P(H_6) = 0.15$, and for simplicity $P(H_0)=P(H_1)=P(H_2)=P(H_3)=P(H_7)=P(H_8)=P(H_9)=P(H_{10}) = 0.0125$, it would be obtained
$P(H_0/E) = 0$
$P(H_1/E) = 0.00035$
$P(H_2/E) = 0.00225$
$P(H_3/E) = 0.00581$
$P(H_4/E) = 0.12185$
$P(H_5/E) = 0.66111$
$P(H_6/E) = 0.18278$
$P(H_7/E) = 0.01357$
$P(H_8/E) = 0.00902$
$P(H_9/E) = 0.00320$
$P(H_{10}/E) = 0$.
And so, the hypothesis largely more probable would result H_5.

I wish to present an even more significant example, for its paradoxical aspects. Paolo does not go to a job interview at the Smile company, we indicate with E this event. The manager of the recruitment department wants to understand why and formulates the following hypotheses:

H_1= Paul found another job

H_2= Paul ended up in prison

H_3= Paul won the lottery or any other reason than the others.

If he were to decide on the basis of the likelihoods, in any case he should have to choose the hypothesis H_2, since H_2 implies E and therefore $P(E/H_2) = 1$.

While if the Bayes theorem is used, the most probable hypothesis may not be H_2. In fact, it suffices that they are $P(E/H_1) = 0.6$, $P(E/H_3) = 0.2$ and $P(H_1) = 0.7$, $P(H_2) = 0.25$, $P(H_3) = 0.05$ to obtain

$P(H_1/E) = 0.618$
$P(H_2/E) = 0.368$
$P(H_3/E) = 0.014$,

from which it would be deduced that the most probable hypothesis is the most reasonable!

Bayes theorem, ultimately, is the architrave of the coherence on which the *updating* of probability assessments rests. Through it one can only reduce *uncertainty*, never eliminate it by coming to definite conclusions. From this point of view, it represents an effective form of *non-linear thinking*.

The *updating* of the probability assessments must therefore follow only one principle: that of coherence; this guarantees to *the observer do not violate the rules* in the attribution of probability: between what he had assessed before (the *initial probabilities*) and what he values after (the *final probabilities*). This is when we refer to *learning inductively from experience*!

3. Conclusions

In any case, from all the examples illustrated, it is possible to see how a set of events can be transformed into a *system of events*, when an observer *brings out* a *coherent* probability assessment for them. As we have seen, in general, of these coherent evaluations there can be more than one, therefore it remains the *responsibility* of the observer to choose the one that he considers to best represent his *state of information*, with respect to the set of events considered!

Consequently, the systemic opening of de Finetti's conception so exerts all its methodological richness and exalts the role of the observer, as the bearer and processor of conscious and precisely responsible choices.

From a systemic point of view, we can say that in this type of inductive logic the interactions between agents (events in this case) must be *coherent*, instead of linear or deductible from each other. In this circumstance, coherence is not something related to the rules of formal logic as in the deduction, for deterministic calculation, but it is relative to the *emergence*.

Coherence, therefore, is not something deterministically calculated and derived: it is designed, learned, experimented and then formalized in the more general construction of models and their simulation, as is also done in the context of *fuzzy logic* (Zadeh et al., 1996).

From this point of view the elements of a system are events.

The observer models the emergent system, which is a configuration of events considered interacting with the probabilities assigned by the observer, as well as by physical interactions.

This approach seems to be necessary for the crucial theoretical role of the observer in emergence processes and their modeling.

Having to do with systems (such as physical, biological and social) considered emerging due to the interactions of the components and assuming objectively that the observer is not part of the system, or is part of it but adopts a logic incompatible with the system considered (that is, assuming a linear logic, thinking of acting in a deterministic space), strategies are used, which although they are not wrong, are at least ineffective.

The approach introduced by de Finetti leads to considering probability systems, starting from single events and arriving at event systems.

In conclusion, it is important to underline how the language and logic of the uncertainty, developed by Bruno de Finetti, play a crucial role in everyday life. Thinking devices that allow us to measure ourselves with *complexity* and produce disciplines that interact systemically, so that concepts, analogies, correspondences and invariants are used consistently.

Tools that until now have allowed us to greatly enrich our knowledge and that if used, as Bruno de Finetti has done not only in the construction of his theory of probabilities but also with regard to his fusionist conception of science and not only, even more they will allow us to increase this enrichment based on overcoming those disciplinary barriers, to which he has always tried to get used to mistrust.

References

Anderson J.R.,1983, *The Architecture of Cognition*. Harvard University Press, Cambridge, MA.

Barrow-Green J.,1996, *Poincaré and the Three Body Problem*. Amer. Math. Soc..

Bruno G., 2002, *Logica dell'incerto e didattica della matematica*, Atti Convegno Nazionale UMI, Loano.

Bruno G., 2004, *Logica dell'incerto: un'esperienza didattica*, Atti Convegno, Insegnare la matematica nella scuola di tutti e di ciascuno, Bari.

Bruno G., 2006, *Logica dell'incerto e didattica*, Periodico di matematiche, serie VIII, vol. 6, n. 3.

Bruno G., 2006, *Uncertainty and the role of the observer*, Proceedings of the 2004 Conference of the Italian Systems Society, Springer 2006 (with G. Minati, A. Trotta).

Bruno G., 2009, *Uncertainty, Coherence, Emergence*. Processes of Emergence of Systems and Systemic Properties, G. Minati, M. Abram, E. Pessa Editors, World Scientific.

Bruno G., 2009, *Apprendere dall'incertezza*. Atti del Primo Congresso AIEMS, Roma.

Bruno G., 2012, *Preserving Uncertainty*, Methods, Models, Simulations and Approaches Towards a General Theory of Change, World Scientific.

Bruno G., 2016, *Bank of Experiences: A Tool to Enhance Creativity, Enterprises and Countries,* Towards a Post-Bertalanffy Systemics, Springer, (whit G. Romiti).

Bruno G., in press, *A need for "Systetics"*, Proceedings of AIRS Congress "Systemics of Incompleteness and Quasi-Systems", Springer.

Corning P., 2002, *The Re-emergence of "Emergence": A Venerable Concept in Search of a Theory*, Complexity 7(6):18-30.

Davydov A.S., 1979, *Bioenergetics and the mechanism of muscle contraction*. International Journal of Quantum Chemistry , 16: 5-17.

de Finetti B., 1937, *La prevision: ses lois logiques, ses sours subjectives*, Annales de l'Institut Poincaré, vol.7, fasc.1, pp.1-68.

de Finetti B., 1974, *Theory of Probability: A Critical Introductory Treatment*, 2 volumes (Translated by A. Machi and A. Smith). Wiley, London.

de Finetti B., 2006, *L'invenzione della verità*, Cortina Editore.

Heisenberg W., 1971, *Physics and Beyond*. Harper & Row, New York.

Klir G.J., Bo Y., 1995, *Fuzzy sets and Fuzzy Logic: Theory and applications*. Prentice Hall, Englewood Cliffs, NJ.

Minati G., Pessa E., 2002, *Emergence in Complex Cognitive, Social and Biological Systems*, Proceedings of the Second Conference of the Italian Systems Society, Kluwer Academic/Plenum Publishers, London.

Minati G., Pessa E., 2002, *Emergence in Complex Cognitive, Social and Biological Systems*. Kluwer Academic/Plenum Publishers, New York.

Peskin E.M., Schroeder D.V., 1995, *An Introduction to Quantum Field Theory*, Westview Press.

Pessa E., 1998, *Emergence, Self-Organization, and Quantum Theory. In Proceedings of the First Italian Conference on Systemics* (G. Minati, ed.), Apogeo scientifica, Milano, Italy.

Pessa E., 2002, *What is emergence? In Emergence in Complex Cognitive*, Social and Biological Systems (G. Minati and E. Pessa, eds.), Kluwer Academic/Plenum Publishers, New York.

Rosen R., 1985, *Anticipatory Systems*, Pergamon, New York.

Scozzafava R., 2002, *Incertezza e Probabilità*, Zanichelli.

von Bertalanffy L., 1968, *General Systems Theory*, George Braziller, New York.

von Foerster H., 2003, *Understanding: Essays on Cybernetics and Cognition*. Springer-Verlag, New York.

Zadeh L.A., Klir G.J., Yuan B., 1996, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh* (Advances in Fuzzy Systems - Applications and Theory Vol. 6). World Scientific Pub Co.

G. & L. E. R., Vol. 22 No. 2, 2018

Simone Enrico Casellina [*] - Giuseppe Pandolfo [†]

PROBABILITY (OF DEFAULT) DOES NOT EXIST

Speech on Simplicity and Comparability of the Risk Measures

Abstract

"The probability: who is it? Before answering [...] it is certainly worthwhile to ask yourself: does probability really exist? And what would it be?" wrote Bruno de Finetti in the item *Probability* of the "Encyclopaedia Einaudi". The probability is our guide in thinking and acting under uncertainty. In banking environments – with Basel 2 – it has acquired an institutional role. The probability (of default) influences the bank's capital, the profitability and the convenience of the business and, at the end, the amount of the credit towards the real economy. If on the one hand much time and care were spent in a progressive refinement of the analytical framework for its numerical evaluation, on the other hand, minor attention was pay to the basic concepts related to the "foundation" of the probability. Many disputes in the international working groups are still spoiled by serious misunderstandings which makes difficult the solution of the technical problems. Still today we see proposals and suggestions – de Finetti would have said *superstitions* – which a rigorous debate has definitively filed. This work is an attempt to repossess – in the simpler style, already from its design, in the form of *Dialogue* – the cornerstones of reasoning (and action)

[*] Bank of Italy, Department of Banking Supervision. Internal Models Validation Division. Via IV Novembre, n. 158, Rome, Italy. Phone: +39-06-4792-4991. *E-mail address*: simoneenrico.casellina@bancaditalia.it. The views expressed do not involve Bank of Italy.
[†] Bank of Italy, Department of Banking Supervision. Internal Models Validation Division. Via IV Novembre, n. 158, Rome, Italy. Phone: +39-06-4792-4990. *E-mail address*: giuseppe.pandolfo@bancaditalia.it (or: giuseppe.pandolfo@ivass.it). The views expressed do not involve Bank of Italy.

under uncertainty, in the assumption of the indissolubility between theory and practice, in the strong belief that many practical problems come from lacking knowledge, or misunderstanding, of deep theoretical concepts.

JEL CLASSIFICATION: B26, C02, G21, G22, G28;

KEYWORDS: Bruno De Finetti, Basel 2, Probability Of Default, Subjective Probability, Coherence, Mathematical Models

1. Introduction

This work was published – in Italian language – as "Quaderno Minerva Bancaria" no. 6, with the Italian title *La probabilità (di default) non esiste. Discorso sopra la comparabilità delle misure di rischio*.

The style of the *Dialogue* was inspired – with temerity, we admit it – to the *Dialogo sopra i due massimi sistemi del mondo*, of Galileo Galilei.

The name "Nottola" wants to remind Minerva's owl.

The name "Chiara" evokes the desire to make simple the risk measurement processes, a feature often requested in the banking environment, although it is tinged with of ingenuity when the discussion moves from the theoretical principles to the methods of implementation.

The phrases of "Bruno" came from to the writings of Bruno de Finetti, with a stylistic freedom justified only by the desire to maximize the effectiveness of the arguments. We have assigned to "Bruno" also some phrases of other authors. However, they are always quoted by de Finetti in his writing as well as they are aligned (in whole or in part) with de Finetti's point of view. In other cases, we have assigned to "Bruno" some our phrases, which are anyway heavily influenced by reading de Finetti's books.

We have decided to loot the writings of de Finetti because we believe that his view on the uncertainty, risks, opinions and accountability should became a mainstream, when people discuss about practical problems of risk management, in bank or in insurance company, as well as from a supervisor's perspective. Furthermore, we have to consider the fundamental behavioural and cognitive aspects that recent experiments have revealed constructively in the past three decades.

We could not say better what had already been perfectly said by an intellect not able to be compared to ours, although de Finetti talk about probability in a wide sense, whereas we use it – every day, with huge difficulties of communication – in a technical field, the measurement of banking risks, on which we have reformulated some of his thoughts. We have the strong belief that the pragmatic writings of de Finetti – for the relevance of the technical contents, coloured with a brilliant phrases, filled with evocative images, if necessary enlivened by a sophisticated irony – are the best antidote against metaphysical tendencies in the discussions about the risk measurement in bank or insurance companies.

It is a standard – in the official publications – signpost in the text the references from which the quotations are resumed. In our case, however, this standard would unduly burden the *Dialogue* and the reader would be overwhelmed by referrals to the bibliography. We therefore prefer simply to list the texts used, leaving the reader the pleasure of discovering the contents and the quotes.

2. A dialogue on risk measurement

There were ladies and gentlemen, who wanted to satisfy their thirst for learning. So they decided to meet, to talk about uncertainty and risks as well as their effects and ways to measure them. After the debts and short compliments, Nottola so began ...

Nottola: *Basel 2* has introduced into the practice of the banking business a powerful math, to make the pricing of the contracts, to evaluate the balance sheet, to fix the capital requirement; it has given an institutional role to probability distributions (on the future); it has reformed the organizational profile of firms and authorities, changing the responsibilities and the day-by-day activities. Along with math and statistics techniques, were required an appropriate IT infrastructure, the legal culture and even some tax considerations. Risk management – with *Basel 2* – has become a general problem of "corporate governance", which imposed to define, or redefine and rebuild, a non-trivial relationship between interdisciplinary and specialization. The enthusiasm of the long design phase, however, was paradoxically replaced by a progressive distrust precisely when the new regulatory framework should have been applied, for the unexpected changes in market scenarios due to the evolution of the financial crisis. What do you think about Basel 2, whereas now many people pay attention to the *Basel 3* and we begin to glimpse a version number four?

Chiara: I want to be fair. The progress is marked by human errors, which have become as many sources of learning. Bankruptcies and the ability to absorb shocks, to use them as a staircase, qualify the rational action of each of us.

Nottola: What are your doubts?

Chiara: With Basel 2, through internal models, we have only the opinions of the risk managers, rather than actual information, reliable, verifiable, and objective, on the soundness of banks. The accuracy of the math in Basel 2 is only an illusion. The underlying math in Basel 2 is in truth mechanistic and reductionist facts. The supervisors have become captives of their own instruments.

Bruno: It is ironic! You are in the prison of the prejudice, but you think that the captives are who look inside through the bars. You mocks Basel 2, for his honesty to analyse the banking risks, and you regret to have no more available a standardized methods, which are equivalent to the adoption of an arbitrary opinion, with the dubious merit to be the opinion of an anonymous Mr. Somebody, that you prefer just to avoid responsibility of a choice. Those empirical *escamotage* for the calculation of capital requirements – those *adhockeries*, *ad hoc* solutions devoid of a single guiding principle – seem to be free to subjective judgments, not because they really avoided them, but simply because they do not nominate them. In this way, you do not restrict them and you leave them the whole field!

Chiara: Basel 2 was an unforgivable optimism momentum for those who, by institutional mission, should always make careful and act prudently: the capital of the bank - with Basel 2 - depends on the probability of default of the counterparties …

Bruno: ... and what else you should expect, if not a linkage between the capital level and the risk of the counterparties, as measured by the probability of default?

Chiara: … with Basel 2, if you let me finish, each bank is free to evaluate the odds as it wishes …

Bruno: … not as the bank wants, but through a model, in order to have the consistency in the evaluation, under to the judgment of the Supervision Authority …

Chiara: … which still does not ensure the same probability of default for the same counterparty, if the evaluation is performed by different banks!

Nottola: Why different banks should assign the same probability to the same counterparty? What is relevant to the probability assessment? The *uniqueness* of the counterparty or the *diversity* of the banks?

Chiara: I do not want to deal pedantic and obscure discussions.

Bruno: Not even me. The verbal acrobatics do not enjoy me and I am bored to besiege castles of words built on the clouds.

Nottola: If we want to develop a reliable reasoning, then we must know exactly what we are discussing. Everyone talk about probability, but nobody is able explain the meaning of the probability in a satisfactory way. I would like to understand the concepts underlying to the probability of default and I would know whether it is reasonable to expect assessments aligned by several banks.

Bruno: Giovanni Vailati – an Italian philosopher and mathematician – argued that the worst way to make sure of the degree of knowledge that an individual has of a concept, is to ask him what the concept is *per se*. I am too pragmatic to worry about the *noumena*. Faced with a concept, whatever it is, I analyse the deep and essential motives that formed the purpose for which that concept was introduced, the reason that explain the intimate reason for its usefulness.

Nottola: I revising my question: which is *the scope* of the probability?

Bruno: The scope of the probability is clarify, to each of us, the own degree of doubt about an uncertain event, depending on the partial information which we have and depending on the algorithm used to process them.

Chiara: So?

Bruno: As consequence, the probability is not *a property of the event* but it is *a relationship between the event and a person* who assesses its likelihood. It does not therefore have an autonomous existence, separable from the person that judges it. The probability does not exist, if not for me, depending on the degree of ignorance in which I am. There is no chance, outside the evaluations that can be done with reflection or thought the instinct. And if you assessed the probabilities with the utmost care, as if they were objective, but at the same time with the greatest sense of responsibility to avoid self-deception to believe that probabilities are really objective, then there is no reason to doubt that my evaluation is wrong, because it does not make any sense outside to me, it has no other purpose than to express what my state of mind is.

Chiara: A *state of mind*? If you want to give to the probability an institutional role, if you want to consign to the probability a crucial task to bank regulation purposes, then it should be an objective probability, otherwise your target is not feasible.

Bruno: You talk about of the probability of default as if it was a physical attribute of the debtor, tangible but not observable. A borrower will have a height, a weight, an income, a civil state, and you believe that he can have in the same sense a probability of default, objective as are height, weight, marital status and income, even though not immediately visible. You see in the probability a sort of dark number, with a physical meaning, or it would be better to say metaphysical, to be placed between the *Planck mass* and

Bohr magneton, a mysterious mechanism which rule the happen or otherwise of an event.

Chiara: The banks hallmark their counterpart's whit the probability of default. Depending on this number, we have different capital requirement, different prices for the borrowers, different accounting results, and as consequence a difference in the balance sheet and in the income statement, which are real and objective.

Bruno: That the accounting figures are true, real and objective, is a exaggerate point of view. There is no truth, unique and unambiguous, about the balance sheet. Given the same material things in the assets and liabilities sides, the truth of a balance sheet constructed for the purpose of liquidation is a different truth from that of the balance sheet built to meet the dividend to be distributed to shareholders; and the truths of the budget of a firm where the shareholders are one or are the members of a family or a stable group of people, or still many changeable and unstable people, are many different truths, but all are fair. If you change the aim of the evaluation, the criteria to assess the firm will change accordingly.

Chiara: This his opinion only creates confusion with respect to our discussion.

Bruno: To be honest, it is an opinion of Luigi Einaudi and I agree with him, because this view is useful also to understand what precautions and what precautions will be able to make propositions with sense, in terms of probability. Affinity is perfect and even goes beyond my intentions. You desire on the comparability of risk-weighted assets, as if the values of the assets, pure and simple, were objective.

Chiara: Dividends paid to shareholders and bonuses received by managers are true and real; they depend on the balance sheet, which we should presume reliable, in turn influenced by the probability of default, … *for which you want to deny all objectivity*! If the probabilities are not objective then what are they?

Bruno: I repeat myself. The probability is a mood and you should include it among the psychological feelings: the probability is an *instinct obscur*, *dont nous ne pouvons nous passer*, in the evocative words of Poincare; it is *the expression of our spirit in a state of imperfect knowledge*, the clear image of De Morgan; it is a *reality of the mind*, in the concise statement of Havvelmo; it is – I say – the opinion of a person about of the plausibility of an event (for him) still uncertain, translated in number and subject to consistency constraints.

Chiara: We are discussing about capital of the banks! No philosophy, please.

Nottola: What are the conclusions of this argument? If Bank X lend 10,000 euros to Miss Truth and evaluates her probability of default to 3%, then it set aside approximately 300 EUR, to hedge against the risk of default. And even if the Bank Y lend 10,000 euros to Miss Truth, but it assesses the probability of default to 1%, then it should set aside only 100 euros?

Chiara: Rather embarrassing, and impossible to justify, if not venturing into a labyrinth of sophistries. Miss Truth is always Miss Truth, one and only, with its probability of default. We can't accept a regulatory scheme incapable of seeing the non-sense of two banks that assign different default probabilities to the same counterpart, with all the dramatic consequences for the supervision activities. It is unexplainable as the Supervision Authority, always oriented to urge the uniformity of judgments in front the same facts – to name one, the concept of *sofferenza* – can live together with a *pot-pourri* of assessments.

Bruno: I read in the official documents that *sofferenza* refers to "counterparties in a state of insolvency, even if not legally ascertained, or in similar situations, regardless of any loss forecasts formulated by the bank". For the *sofferenza* the dreaded event has occurred, the risk of default no longer exists, so the probability of default no longer makes sense, and the purpose of the *sofferenzea rettificata* is to make the information available to those who still are not aware about them. The situation is quite different, when we talk of uncertainty and probability.

Nottola: Even the probability depends on the information...

Bruno: ... with a clarification, though. To speak properly about probability you have to know something and ignore something else, because if you know all, then the uncertainty vanish and the probability collapse to 0 or 1, the only true values in a strict sense, referring to two extreme and pleonastic cases, in which the idea of probability is no longer useful. If we are not ignorant, then there no probability! The concept was already clarified by Jacob Bernoulli: *probabilitas est gradus certitudinis et ab hac differet ut pars in toto*, to suggest a quantitative assessment of knowledge in the psychological premise of their limits. Then again Keynes: every proposition is true or false, but our knowledge in this regard depends on the circumstances.

Nottola: What do you want to proof?

Bruno: The judgment of the probability of the same event can change from one person to another – from one bank to another, if we are interested to credit risk – because are different the respective partial information.

Chiara: We could imagine a set of information, which although incomplete, is favourite with respect to the assessment of the probability.

Bruno: The only privilege would be known all, do not have more uncertainties, and therefore make less of the probability. But in every concrete situation of uncertainty, the degree of informational incompleteness is a matter of the problem and we must learn to live with it. You may decide to acquire new information and then you will have a new probability, that does not contradict the previous probability, but it simply passes from one thing (the probability in a certain state of information) to another (the probability in another state of information) which is different from the previous one to be in agreement with it. If I say "my watch marks 11.25", I do not correct the assertion made an hour ago, "my watch marks 11.25", but I modify it to preserve the coherence.

Chiara: We could at least ensure that all banks have the same set of information, in assessing the probability of default.

Bruno: The banking business is based on private information, the imbalances of knowledge are unavoidable and they qualify advantages and disadvantages in the market. No bank would be willing to share its information set. On the other hand, if you want to share common information and then oblige banks to make a probability assessment subordinated just common knowledge – a sort of intersection of information available in the banking system – you would get to the undesirable situation of a measure calibrated on the smallest set of information. However, even by omitting business considerations – at the end I am neither an economist nor a banker – there is a discourse that I care most.

Nottola: What?

Bruno: A hypothetical coincidence of set of information, if can facilitate the uniform judgments, does not guarantee it The information sharing does not bind to a convergence of *views,* because everyone is still free to be understood in its own way the different parts of the same body of knowledge It would be a blessing, if when we talk about the probability of an event *E,* we keep in mind the writing $\mathbf{P}_{\mathrm{I}}(E|\mathrm{H})$, where is clear the dependence of the probability **P** from the set of information H as well as from the individual I that judges, instead the simplistic notation $\mathbf{P}(E)$, which brings up the probabilities as something eternal and immutable, like the smile of the Cheshire Cat, which remains even after the cat went away.

Nottola: We could establish at least fix a boundary, an upper and a lower limit, a range within which a probability assessment must fall, to be reasonable.

Bruno: I do not agree, because, instead of deleting the embarrassment of choosing a precise value for the probability, doubles it, forcing you to choose two precise boundaries.

Chiara: If everything can be true, when we talk about risk-measurement, then this confirms the need to rethink the regulation. A regulation framework based on opinions does not work. It serves objectivity, at times even conventional parameters, but based set on objective facts.

Bruno: The importance of knowing the facts, to know them in the form of numerical statistical data, to have a mentality in which they occupy a crucial position, is out of the question. *But a fact is like a sack: blank does not hold*. probability with the facts as you do a house with stones, but a bunch of facts is not a probability like a bunch of stones is not a home. You can gather factual data as judgments, but the answer is not in fact. It remains in subjective *opinion,* that no fact can constrain, but that certain facts may possibly spontaneously and feel constrained.

Chiara: I cannot accept this point of view …

Bruno: Why? Every day we can see paradoxical conflicts, in which, in support of opposing arguments, the parties invoke the same facts – on which they are agree, sometimes even in the smallest details – but from which they draw different conclusions.

Chiara: Please, we are not discussing about ordinary everyday problems, where everyone is free to decide. Here we talk of the capital of banks, although the frequent digressions seem to aim at confusing.

Bruno: No digression from my side. I just do not believe in a way of thinking for the day-to-day decisions and in a different way for the technical, scientific, or institutional issues. I think, by contrast, that there is simply a unified concept for the things of the world, very good for all applications that can be the bets on sport events or the decisions about capital levels of the banking systems.

Nottola: But *in practice* what should we do when faced with a discrepancy of probabilistic assessments of the same event?

Bruno: Knowing that different banks set a different value for the probability of one event is a fact that should be used, but not to mediate between evaluations, hoping in the miraculous *in medium stat virtus*, or to force the lowest value to the highest, in an illusory exercise of conservativeness. We must use this fact to better understand the problem, in terms of information and models used.

Nottola: We see an immense task for supervisors. Are we sure that they are capable?

Bruno: I do not know, but *people had no choice*, as Robert Merton commented on the inevitability of the use of Black-Scholes formula, although it was made of esoteric concepts, outside to the curricula of market participants, even if then they began to speak of hedge ratio and delta, to appoint the lemma Itô and differential equations.

Chiara: Philosophy, philosophy! We will see if it is not possible to give the probability an objective meaning. When you bet on the launch of a coin or a dice, when you play roulette or betting, you have an objective probability of win. In switching from the case of the games, *per se* trivial, to more relevant uncertainties, there is no reason to think that there are not the same objective probabilities.

Bruno: You must keep me apologized, but *I do not understand*: what does it mean that the probability of "head and cross" is objectively the 50%?

Chiara: I do not care about coin. I quoted the paradigm "head and cross" to describe a situation where you have a true probability, 50 and 50, and I do not see why it should not be so elsewhere.

Bruno: The idea of a true probability is an aberrant superstition and it is most disappointing thing to hear again - presented as critical - so superficial clichés. I do not understand if who believes in certain misconceptions has read the arguments to refute them or if they have read them without understanding anything or instead if they have understood them in reverse.

Chiara: If you want to make it clear also for us …

Bruno: You talk about of probability of the coin as if it were the probability inside the coin, melted in the coin in a physical sense, in one with metal alloys from which it is composed, as well as you speak of the probability of default as if it were a genetic item of the counterparty. It seems to me to do the utmost to understand the discussions of others, to respond carefully and patiently, even when it comes to repeating things that have been said and repeated, in order to argue trivial misunderstandings. Rarely I have the pleasure of having the impression that others make a similar effort.

Nottola: I invite you to make an effort again …

Bruno: The coin has a shape, has a diameter, has a thickness, has a weight, but it cannot be said, in the same way, which has a probability. The Polygraph, among the technical specifications of the mintage, certainly does not take into account the probability of the two faces. The probability is not – it cannot be, even to want – a physical characteristic of the coin. Many people love to say “perfect coin”, in their discussion on the probability, but nobody can formalize this concept.

Chiara: But in the case of the coin, all agree in assessing the probability 50 and 50.

Bruno: Because everyone is in the same state of mind. I estimate the probability to 50% because I feel a symmetrical uncertainty respect to the two faces, because I cannot find reasons to give priority to the output of one on the other face, because to be perfect or symmetrical is my ignorance in front of the outcome of the launch, not the coin itself. Then, the fact of experiencing a universally shared mood, the fact that everyone evaluates in 50 and 50 the probability of "head and cross" does not transform it into something different from my mood, does not make it objective simply because unanimous. If you think to reach objective conclusions, by layering the subjective opinions concordant, is like believing that a heap of stones, increasing more and more, at the end will eventually become an animal.

Chiara: When you extract the balls by urn - and forgive me if I insist on school cases, but it is only to show clearly the general idea - when you extract the balls, I said, the chance of having a red or a yellow ball is objectively determined by their proportion within the urn.

Nottola: So, if the urn contains 50 red and 50 yellow balls, the probability is objectively 50% for each colour and if you believe in a different probability, then you are wrong.

Chiara: Yes. Of course, the urn is merely a trick, but it broadly defines a general schema, useful to understand the problems, in the areas of actual interest.

Bruno: Imagine then having to assign a judgment to the probabilities evaluated by three individuals about the extractions from an urn, in the same way that a supervisor is called upon to express himself or herself on the quality of the probability of default assessed by different banks. The first individual declares a probability of 50%. Is an acceptable judge?

Chiara: It is the only acceptable one. An alternative assessment would be extravagant, incomprehensible, without justification.

Bruno: The second individual, however, declares a 40% probability for the exit of a red ball.

Chiara: Unacceptable!

Bruno: Why?

Chiara: The urn contains red and yellow balls in equal proportion, 50 and 50, randomly shuffled, so …

Bruno: I know, I know, …. Now we are addressing the third person, who says with confidence 99%.

Chiara: A crazy!

Bruno: However, we are curious to know his justification. Therefore, we approach him to ask explanations and we realize that from its position, from his viewing angle, from his perspective, thanks to a game of mirrors, he able to discern clearly the urn content. Now we also see what he sees. The random shuffling of the balls has given rise to a rather curious arrangement: all fifty reds up, the fifty yellow down. Is his 99% probability still a madness?

Chiara: Not anymore.

Bruno: Now is it the first individual who stated 50% to make a mistake?

Chiara: Yes!

Bruno: Please, let us emend him!

Nottola: … but we should give to him the private information of third party; otherwise, he will not agree to deviate from the most natural evaluation.

Bruno: Please, no scruples of confidentiality, faced with the possibility to know and spread the true probability! We surely go to the first person, whispering in his ear the particular arrangement of balls known to the third.

Nottola: Now the assessments will be the same …

Bruno: … no, unfortunately, because once extended its set of knowledge, the first individual with equal confidence answers 1%!

Nottola: I do not understand …

Bruno: Simply, he knows the typical behaviour of who extract the balls, to have repeatedly observed him in the past. Almost never he extract in surface. He loves instead sink the urn arm, in depth. As consequence, there are low chance to have a red ball. What do we think of now of this probability?

Chiara: This is the right evaluation …

Bruno: Well, depending on the available of new information, it was baptized “true” a probability of 50%, of 99%, and of 1% and we don't care to hear the reasons of the second individual, which had evaluated the probability at 40%, which may have similar arguments as well, if not more convincing than the other two, in support of its assessment. Each of us evaluate the probability with the information that he has, for each of us there is a personal probability, as Savage calls it, each in his own way, as Pirandello said, each with its own probability, no more true or false than the other, which only require consistency with other probabilities evaluated by the same individual.

Nottola: I am more confused than convinced …

Bruno: The confusion arises because people confuse the composition of the urn, which is a physical fact, with the probability that is a psychological state dependent on information. The examples, if desired, may multiply. Facing physical states in appearance symmetric - the two sides of a coin, the six faces of the dice, the thirty-seven chance of roulette or the composition of an urn - there will always be additional information that, if known, will influence our opinion and all the physical symmetries will be destroyed with their associated probabilities.

Nottola: So?

Bruno: So, the objective probability never exist , even in the very special field of play - where we have only a convergence of opinions, due by particular conditions - and a fortiori it cannot exist in the field of banking risks, dominated by information disparities and divergence of beliefs. Each bank will have its model, built on judgments, information and motivation you will be able to know or not to know, to share or not share; you can judging it more or less reasonable - with double subjective judgment: our in comparison to the other -, and it can be more or less close to those of a few or many or all other banks. But in an evaluation of probability we can judge only if it is or is not consistent , and consistency is ensured by its use of a mathematical model, whereas if you would like evaluate the probability only with the intuition or instinct, then it's high the danger to fall in irreconcilable contradictions.

Chiara: I will not leave you a tooth that I consider healthy. If you launch a coin a high number of times, you will have roughly the same relative frequency of heads and crosses, and the goodness of the approximation increase with the number of the launches. The probability is objectively 50% and by analogy, out of the case of games, you can still think to determine it experimentally, through repeated observations of similar cases.

Bruno : *Probability* and *Frequency* , the female version of the Gemini of Plauto …

Nottola: The Gemini of Plauto?

Bruno: Yes, *The Two Menecmi*, a comedy of Plauto, where two Gemini are continuously exchanged between them, arousing continuous, funny and absurd misunderstandings.

Chiara: What is the connection with our discussion?

Bruno: There are so many and complex connections between probability and frequency, two notions each other extraneous, that if you confused the one for the other, then any attempt at clarification will make the situation worse.

Chiara: It is all clear, from my point of view. The probability is an idealized frequency, and I am not afraid to say mythologized: it is the frequency that you should have in an infinite number of experiments, assuming they can be run in the same conditions.

Bruno: You present yourself as a pragmatic decision maker, but in your discussions appears the infinity, I mean situations for which no one know what will happen. In the long run ... we're all dead, even the survivors of the famous long period of Keynes. And then what does it mean repetitions performed in the *same conditions*? If the repeats were seriously carried out under the same conditions, then they would give invariably the same result, right? This confusion between probability and frequency – between probability of default and default rate, we can say – is a bad mistake in the teaching of probability theory and its persistence is paradoxically due to the fact that it is the worst interpretation of probability.

Chiara: The insurance companies thrive on the *frequency-probability*, their entire business is built on the idea of a frequency that approximates at the best to the probability. Insurers behave as if the frequencies were probabilities, and with these *frequencies-probability* they calculate the premiums, set the technical provisions, and fix the own funds as well as pay the dividends. If things work out, thanks to the *probability-frequencies*, cannot be just their feeling, we must have under an objective reality.

Bruno: Well said! It is *as if*, it is the *als ob* of Vaihinger. Why if something works in practice, do you want to see the philosophical substratum of an absolute truth?

Chiara: The frequencies are objective …

Bruno: I do not agree. To perform the frequency calculation you have to group the single observations into the clusters, because only inside a class you can count the "successes", to relate to the total, to have frequency; and grouping criteria are subjective , they only have practical value and no logical necessity; any attempt to use them to have an objective definition of probability is illusory. Two objects can always bring together in the same class or separate into distinct classes, and among the many classes there will be some subjectively interesting, for which it will judge useful to introduce a special name. But it is only a question of utility and any philosophical discussion would be trivial. There is no need to seek the truth, we asks only to become aware of own opinion.

Chiara: The probability of death depends objectively on the age of the insured, or do you want to deny it?

Bruno: I do not deny it, although I would prefer to speak of broad inter-subjective agreement. However, I admit of course a strong dependence – and I can say objective, if it can give pleasure – between age and the risk of death. So, given a population of N individuals of age x, of which n are dead, can we say that n/N is the objective probability?

Chiara: If the number N of the initial population is high enough …

Bruno: Let us imagine that it is. Then n/N is the best estimate for the unknown probability?

Nottola: Not really. Age is a critical risk-driver, but the statistics says also that the life expectation of women is higher than men's, so the N units should be classified by age and sex.

Bruno: And when we have grouped by age and sex, at least we got *the best estimate of the true unknown probability*, as someone likes to say, with an expression of which you should suspect only for its length.

Nottola: Not really. There are so many other factors to consider, if we want to be precise.

Bruno: We want it …

Nottola: The nationality, if necessary qualified for the region where you live; the marital status and eating habits; it's relevant to know if you practice or not some sports, unless it is extreme sports, which would be an aggravating factor; indeed, now that I reflect on this point, even ordinary sports increase the risk, If you did not have the doctor's permission before; there is also the attitude to avoid or not some dangers, I think for example to the driving style. All these factors intuitively linked to life expectation.

Bruno: We do not neglect the results of a check-up, often required to stipulate the insurance contract.

Nottola: Right!

Bruno: I would like to note that the *objective* risk factors, objectively to take into account to have an *objective* valuation of the *objective* probability, have the disadvantage to empty the groups to calculate the frequencies. If the initial population of age *x* count N=100.000 individuals, distinguishing by sex you could have for example N_M=40.000 males and N_F=60.000 females. If we add marital status, the subpopulation of the males could split in N_{MS}=15.000 married and N_{MC}=25.000 bachelors, and instead for female we could have N_{FS} = 35.000 married and N_{FN} = 25.000 unmarried. If we introduce other risk factors, then our groups will become gradually poor, with few observations to calculate the frequency, and as consequence our empirical estimation will be unsatisfactory.

Nottola: I glimpse a merciless conclusion …

Bruno: If you are hunting the ghost of the real probability, if you chase the Morgana Fairy of the objective probability, then you will end up accommodating so many risk factors that the observation will remain only one, no statistics will be possible and you will be forced to a personal evaluation. To get the highest possible level of objectivity, you will discover to be subjectivists!

Chiara: You had to stop before, of course. You have to stop the classification process *at the right point*.

Bruno: Who decides that? On what basis? Why should be included some risk factors and other not?

Chiara: It is necessary to take into account *the narrower class for which we are still able to determine the probability*.

Bruno: There climbs on the words, in a desperate tentative to give meaning to what has none. We passed one after the other the objective probability, the frequency, the right point of arrest and narrower class. At each request to clarify a concept from an operational point of view, you propose some other indefinite concepts, in the tacit hope that from a stream of words will emerge, by parthenogenesis, a logical proposition or a mathematical formula or even just a rule of thumb, or simply that your counterparty will decide to abandons the discussion, for exhaustion. The difficulties to objectify the probability, as opposed to the Guard of Napoleon, always retreat, but they never die. Each definition apparently objective requires subjective judgments, which can be hidden, not deleted. Of course, you can make all perfectly logical, if you remain in silence about your judgments, you can renamed them as hypothesis, but the value of these assumptions depended on personal judgments, and it is well that we must talk of these as essential fact.

Nottola: Also the non-life insurance companies are in the same condition?

Bruno: Obvious. If you want to say "the annual frequency of the car accidents is 5%" – and then you want to use 5% as probability, to conjecture the number of the future incidents - will be necessary to bring together the individual drivers into classes, so that you can count the number of claims, to have the frequency. In which way will you build your classes? By city? By age of the driver? By sex? By number of past incidents? By type of the car? Or through a combination of some or all of these and other variables? Each class will have its own frequency - so its probability, admitted to want to assimilate frequency and probability - so each driver will have one, none and hundred thousand probabilities to have an accident, as many as the possible classes where you can place him.

Nottola: Can we apply the same point of view to the banks?

Bruno: It cannot be otherwise and I would like to expand the perspective of our analysis. I read from a book - written over thirty years ago and its author would then become a central banker - and I want to emphasize some passages: "A mathematical model, even in its most refined details, can provide support for the decision within limits. These limits exclude that the decision runs arbitrariness. Beyond those limits, the decision is left to the appreciation of discretionary individual cases based on the banker's experience and judgment. The final choice is therefore necessarily subjective. The probabilities of various outcomes of the loan, taking into account all the information that is useful acquire about the company to be given, shall be assessed only subjectively. Even when we rely on a mathematical model, an art element cannot be eliminated, at least in the estimations for the values of the exogenous variables on which each model is based. As consequence, the same loan, convenient for a bank, may look to other unacceptable for other banks for several reasons: different assessment of the probabilities of the individual outcomes, different composition of the credit portfolio, the different degree of risk aversion, different ability to deal with the risk"

Nottola: … *subjective choices*, … *subjective evaluations*, … *unavoidable art elements*, … *different probability assessments*, … I do not understand. The mathematical models, with all their paraphernalia of formulas, data and calculations, have the aim to ensure objectivity in the risk measurement and standardize the decision processes. Or not?

Bruno: The outcome of a model is a number, not a decision. Models do not objectify anything. Their aim is rather to declare, clarify, explain the unavoidable elements of subjectivity, required to reason and decide under uncertainty conditions. They ensure consistency in the evaluation and make them transparent to third parties, if technically equipped. With a model is easier to formalize a coherent opinion and realize a comparison between different viewpoints.

Nottola: The most recent proposals of the Basel Committee on the regulation of internal rating systems, however, seem to mark a deep difference between two kinds of situations. On one side, we have the situations for which there is a huge of statistic information, and on the other side, we face instead situations where the empirical evidence is almost absent. We can think at the contraposition between high default and low default portfolios. For the residential mortgages, for example, we have always a well-defined set of historical default rates, high or low depending on the business cycle, and we can perform a reliable estimation of the probability. However, for other portfolios, for example large corporate, banks and sovereign, the default event is an "isolated" event, sometimes completely absent, and then no risk estimation is possible.

Chiara: Exactly! This distinction – between high default and low default – is related to the contraposition between risk and uncertainty. We must separate the concept of risk from the uncertainty. We must distinguish between the peculiar uncertain situations where there is a large agreement of the opinions and therefore easier possibilities of measurement and control, and the situations for which the lack of shared information does not permit a numerical evaluation, and so we can only undergo or avoid the effects, or face them with a purely conventional style.

Bruno: This discussion about risk and uncertainty – initiated by Keynes and Knight in the 30s of last century – can now be considered resolved in the foundations of the game theory. A criterion of well-defined decision should have general validity, not being deducted from specific empirical hypotheses, but from general and simple logical consistency conditions. It makes no sense to restrict the validity of the criterion to cases of "risk" based on assumptions unnecessary and ill-defined, leaving the other cases without a criterion or with an arbitrary number of more or less artificial solutions. Moreover, if you define as "risky" the situations involving minor differences from one individual to another, your definition is weak, because for judgments is always a matter of degree. Therefore, the distinction between risk and uncertainty, in the sense proposed by Keynes and Knight, is not expressive. The greatest danger is to make her feel net and fundamental instead nuanced and ancillary.

Chiara: The contrast between risk and uncertainty is aligned with the intuition. We feel indeed a difference between a bet on "head and cross" and a bet on a football match and we require different odds, to participate at the game; in the guise of an insurer, we will fix different contractual clauses to take in our portfolio the risk of a car accident or an earthquake.

Bruno: No doubt. These moods are rooted in common feeling and I do not want to put them into question. I say only to pay attention to the danger of misunderstandings. You make a bad mistake, if you believe that the differences in uncertain situations have the power to limit the scope of the probability theory, whereas they are only linked to the outward items, which could have a role to address the evaluations and behaviour of each us, without affecting the validity of probability theory, its universal applicability. The situations of uncertainty differ only in degree, not by nature, and the probability theory is a unique and general method to rule the uncertainty, it is a unitary theory that welcomes in a general framework the different partial views, without forcing them.

Chiara: I'm not convinced at all. When you have a large number of observations - even if they were merged into classes in a subjective way, even if they were organized into groups based on personal judgments – we are in an environment where a mass effect works. Otherwise, without a pile of observations, we have only individual effects and of course we can't rule them. Thus, the measurement strictly depends on the existence of a pile: as long as you have a few observations, they do not constitute a heap and we can conclude nothing, but if there is a pile, then, but only then, we have some hope to perform a reasonable measure.

Bruno: Let me understand: if you plan on adding a remark at a time, nothing could be said until their number is insufficient to form a pile, but the conclusion will leap out, finally, when the *no-pile* will be turned into a pile. Suddenly? Going from 99 to 100? O from 999 to 1.000?

Chiara: This is a caricature. Of course there is not a sharp jump …

Bruno: … of course! The no-pile passes through a to-be-or-not-to-be-an-pile phase, inclining first one way and then the other, and only subsequently does it gradually transform itself into a real and genuine aggregate …

Nottola: Even more grotesque and vaguely ironic …

Bruno: Inevitably! Because we do not answer the original objection raised against the distinction, here put forward as being of fundamental conceptual importance, between the pile effect and the effect of individual elements.

Nottola: How do we come out?

Bruno: We must deny any such distinction. The conclusion which can be reached on the basis of a mass of data is determined not globally, as a mass effect, but as cumulative effect of the contribution of every single information. I do not want to hide the differences between situations of uncertainty objectively different. Simply I do not want to dramatize them, creating an artificial classification, which leads towards inadequate and distorted interpretations. In general, the relevant information required to fix an opinion could be different – from individual to individual, from bank to bank – and depending on the circumstances the opinions – the probabilities – will have different degrees of stability and sharing. Some of them will be well established and widely accepted, other will be more undulating, others will remain interim, but they are all opinions and they remain opinions. We can also make a distinctions – for example high default versus low default – if we think that they are useful, but these distinctions concern only external aspects of the available information, and do not affect the unitary nature of the probability and its application in all fields. On the other hand, when a banker lends money, when he makes a price for the loan, when he draws up the financial statement or makes a decision about capital level, he is assigning a probability, even if he is not aware of it, and the banker must do all this regardless if his counterparty is into a high or low default portfolio. The refuse to evaluate a probability, on the pretext that there is not enough empirical evidence, it means acting like Cremonini, scientist contemporary of Galileo, which, with some folklore, it is said he did not want to look through the telescope to avoid to change his beliefs about the cosmos.

Nottola: Excuse me, but in all that, what happened to the axiomatic theory, which gives a *mathematical definition* of probability?

Bruno: The axioms and theorems of the probability theory fix only a huge of constraints between individual evaluations, totally free in all other respects. Given two events A and B – prepositions for which we will able to say if they will are true or false – a classic theorem of probability theory is written in the form:

$$\mathbf{P}_{\mathrm{I}} = (\mathrm{A} \cup \mathrm{B}|\mathbf{H}) = \mathbf{P}_{\mathrm{I}}(\mathrm{A}|\mathbf{H}) + \mathbf{P}_{\mathrm{I}}(B|\mathbf{H}) - \mathbf{P}_{\mathrm{I}}(\mathrm{A} \cap \mathrm{B}|\mathbf{H})$$

This means: "the individual I, with his information set H, must assess the probability of the occurrence of at least one event between *A* and *B* so that its assessment equals the sum of the probabilities of the individual events, minus the probability to have *A* and *B* together". Mathematical probability theory shields your opinion from the contradictions as well as the geometry ensures the consistency in the measures of the physical objects. If anyone, after measuring a rectangle, said he had found a base of 5 cm, a height of 10 cm and an area of 25 cm^2, we can say to him to revise at least one measurement, without discussion on which measure, because it is not true that 5×10=25. The axiomatic theory only state the constraint $\mathbf{P}_\mathrm{I} = (\mathrm{A} \cup \mathrm{B}|\mathbf{H}) = \mathbf{P}_\mathrm{I}(\mathrm{A}|\mathbf{H}) + \mathbf{P}_\mathrm{I}(B|\mathbf{H}) - \mathbf{P}_\mathrm{I}(\mathrm{A} \cap \mathrm{B}|\mathbf{H})$, but it does not give you an instrument to evaluate the single pieces of the formula. Therefore, we have an infinity of assessments itself coherent, because they respect the constraint of consistency. Everyone will choose what he prefers or, to be right, one will choose what he feels. In conclusion, the models are useful, applicable everywhere and appropriately, provided that each one always start from his opinions; but the models become illusory, and therefore dangerous, if you plan to use them in a miraculous way, to avoid to declare your view, thinking to set a tool to create the opinions, whereas it can only to manipulate them.

Nottola: To return on the credit risk of a counterparty …

Bruno: … the same loan is otherwise risky if placed in a different portfolio, if it is considered from different banks.

Nottola: I already see a theatrical grimace of disappointment, on the face of regulators and supervisors.

Bruno: I am sorry, but there is no way in which the individual can avoid the burden of responsibility for his own evaluations. The key cannot be found that will unlock the enchanted garden wherein, among the fairy rings and the shrubs of magic wands, beneath the trees laden with monads and noumena, blossom forth the flowers of *Probabilitas realis*. With these fabulous blooms safely in our button-holes we would be spared the necessity of forming opinions, and the heavy loads we bear upon our necks would be rendered superfluous once and for all.

Chiara: However, I continue to see a link between probability and frequency similar to the linkage between the true measure of a physical quantity and its experimental evaluations affected by errors due to the measuring instrument.

Bruno: The analogy is pushed beyond the reasonable, because forgets – to make a concession of jargon and mind-set – that the probability is an atypical measure of an invisible object. The reliability of a frequency for the evaluation of a probability depends on your beliefs. It make no sense to say "probability" instead of "frequency", thinking to make objective the probability, if the connection between the technicalities to detect the frequency (in the past) and the way to have an opinion (about the future) remains subjective. It is better analyse directly the subjective element in which is rooted the notion of probability.

Nottola: Anyway, there is a strong psychological reluctance to abandon the idea of objective probability.

Bruno: We are unfortunately under the *tyranny of language*. The word *subjective* evokes the most bizarre and capricious attitudes, whereas objective would means a serious and reasonable behaviour. The hesitation to accept the subjective probability, in the sciences and applications, is due to the impression that if everything is subjective, then everything has to be arbitrary and no rule can enforce. But in its technical meaning, free from emotional influences, subjective simply means *related to a subject* and subjective assessments are estimates made by a person, by an individual, who takes responsibility and accepts the consequences. Subjective is all that requires intervention from the judgment of an individual, because you cannot reduce it to the mere finding of a fact. The views of a mathematician and of a cabbalist, of an astronomer and of an astrologer, of a doctor and of a self-styled healer, are certainly subjective, equally subjective because all expressed by individuals, and you could grant them more or less confidence. The subjective probabilities school does not want to accept superstitious or metaphysical interpretations, only it keen to emphasize the need to recognize, in all reasoning, what is logical, what has an empirical meaning and what is a pure subjective value, because this distinction should be established in any mathematical theory to be able to usefully deepen criticism of principles.

Nottola: It is a revolutionary change!

Bruno: I understand very well the difficulties that those who have been brought up on the objectivistic conceptions meet in escaping from them. I understand it because I myself was perplexed for quite a while some time ago. It was only after having analysed and mulled over the objectivistic conceptions in all possible ways that I arrived, instead, at the firm conviction that they were all irredeemably illusory. It was only after having gone over the finer details and developed, to an extent, the subjectivist conception, assuring myself that it accounted for everything that is usually accredited, over hastily, to the fruit of the objectivistic conception, it was only after this difficult and deep work, that I convinced myself, and everything became clear to me. It is certainly possible that these conclusions are wrong; in any case they are undoubtedly open to discussion, and I'm glad we discussed it.

Chiara: However, my point of view is paradoxically strengthened precisely by the arguments of my counterparty. If he proclaims the subjectivity of probability of default, then I do not understand how you can think to build a banking regulation on this basis. The rules would be based on the *soft mud of the feelings*, the entire institutional architecture would be built on *the sand of the internal models*.

Bruno: If you remove the sand, then you will build on the empty. The regret of losing faith in the objectivity of probabilities, and in the regulatory measures that use it, it is unjustified. Nothing is lost, except what was an illusion. However, because everything in life is uncertain, and outside of that, nothing can assert with certainty, I do not want to preclude the possibility to change my idea. Therefore, if a day you you'll have access to the strongbox of the Platonic Truths, and you will explore it to see if there are or not the *true probabilities* that you are looking for, I promise to adopt them, if you will find them. Until then, I will not care about of the objective truths, and I will work with something formally much less, but substantially much more: *my opinion*.

Nottola: We live in a complex world. The financial system has long globalized and now also the regulation and supervision begin to be globalized, overlapping and mixing different experiences, traditions and cultures. How can we orient ourselves in such a multifaceted world?

Bruno: The North Star is the commitment to be honest and earnest, against *the hundred ways to say nothing*; it is the engage to have and declare the *opinion*; it is the constraint to be *coherent*.

3. Conclusion and next steps

This *Dialogue* wished to put the conceptual basis for a consistent discussion about risk measurement in banking and insurance environments. We will be now committed to translate the main theses of the *Dialogue* in a technical content. The main issue for regulators and supervisors of the financial sector is double-faced: on the one hand, they want to understand how much confidence can be granted to the risk measures proposed by the intermediaries; on the other, they want to be sure that firms have the most effective incentives for a measurement that, without pretending to be "true", should be at least as earnest and honest as possible. The binomial test – when we treat about credit risk – is methodologically weak, with respect to these requirements. Our proposal, on the first stage, use a Bayesian "credibility" mechanism to mix the views of the regulator and the company into a single probability measure (of default) to be used to calculate capital requirements. Then we propose the use of the penalty method – "the only meaningful way to deal with the prediction problem", to quote de Finetti – to assess the probability measures proposed by banks and push them towards the *best-effort* to have their *best-estimate*. An appropriate calibration of our proposal allow to rule in a general scheme the so-called *Trought-the-Cycle* and *Point-in-Time* probabilities, to meet effectively the requirements of the supervisors and accounting standard setter.

References

de Finetti B. (1957), *L'informazione, il ragionamento, l'inconscio nei rapporti con la prevision*e, in: "L'Industria", n. 2.

de Finetti B. (1965), *La probabilità: guida nel pensare e nell'agire*, Trento, Quaderno n. 11 dell'Istituto Universitario di Scienze Sociali di Trento.

de Finetti B. (1967), *L'incertezza nell'Economia*, in B. de Finetti, F. Emanuelli Economia delle Assicurazioni, vol. XVI del Trattato Italiano di Economia, Torino, UTET.

de Finetti B. (1970), *Teoria delle Probabilità. Sintesi introduttiva con appendice critica*; (2005), Milano, Giuffrè, (ristampa anastatica).

de Finetti B. (1971), *Probabilità di una teoria e probabilità dei fatti*, in "Studi di probabilità, statistica e ricerca operativa in onore di Giuseppe Pompilj", Gubbio, Oderisi.

de Finetti B. (1980), *Probabilità, in Enciclopedia X: Opinione-Probabilità*, Torino, Einaudi.

de Finetti B. (1981), Scritti *1926-1930*, Padova, Cedam.

de Finetti B. (1989), *La logica dell'incerto*, Mondadori M., Milano, Il Saggiatore.

de Finetti B. (1991), *Scritti 1931-1936*, Bologna, Pitagora Editrice.

de Finetti B. (1993), *Probabilità e Induzione*, by Monari P. e Cocchi D., Bologna, CLUEB.

de Finetti B. (1995), *Filosofia della probabilità*, by Mura A., Milano, Il Saggiatore (p. 71-72).

de Finetti B., Savage L.J.(1962), *Sul modo di scegliere le probabilità iniziali*, in "Biblioteca del Metron", Istituto di Statistica dell'Università di Roma.

We have also consulted the following publications:

Banca d'Italia (2008), *Matrice dei Conti*, Circolare n. 272 del 30 luglio 2008, aggiornata al 20 gennaio 2015 (p. B.6).

Barra M. (2001), *Sulla «legge dei grandi numeri», sulla «legge empirica del caso» e sull'apprendimento attraverso l'esperienza, anche in connessione con alcune questioni del «teorema del limite centrale»: aspetti storici, epistemologici e didattici*, Convegno Internuclei per la scuola dell'obbligo di Monticelli Terme-PR, 5-7 aprile, available to the link www.brunodefinetti.it/Bibliografia/Barra.pdf

Bruno G., Giorello G. (2006), *Scienza senza illusioni, in de Finetti B., L'invenzione della verità*, Milano, Raffaello Cortina.

Cifarelli D.M., Muliere P. (1989), *Statistica bayesiana. Appunti ad uso degli studenti*, Pavia, Iuculano

Ciocca P. (1991), *Banca, Finanza, Mercato, Torino*, Einaudi.

Ciocca P. (2000), *La nuova finanza in Italia. Una difficile metamorfosi*, Torino Bollati Boringhieri.

Einaudi L. (1940), *Miti e paradossi della giustizia tributaria*, Torino, Giulio Einaudi Editore (pp. 183-184).

Daboni L., Wedlin A. (1982), *Statistica. Un'introduzione all'impostazione neo-bayesiana*, Torino, UTET.

De Felice M. (1999), *Le circolari della vigilanza e la cultura bancaria*, Roma, "Surplus", Vol. 1, n. 1.

De Felice M., Giorello G., Moriconi F., Piccinato L., Salinetti G. (2014), *Conoscere de Finetti per il governo dell'incertezza*, Milano, Mondadori.

De Felice M., Moriconi F. (1991), *La teoria dell'immunizzazione finanziaria*. Modelli e strategie, Bologna, Il Mulino, pp. 17-18.

de Finetti F., Nicotra L. (2008), *Bruno de Finetti, un matematico scomodo*, Livorno, Belforte.

Lindley D.V. (1990), *La logica della decisione*, Bologna, Il Saggiatore.

Keynes J.M. (1994), *Trattato sulla probabilità*, by Pasquinelli A., Marzetti Dall'Aste Brandolini S., Bologna, CLUEB.

Piccinato L. (2009), *Metodi per le decisioni statistiche*, Milano, Springer.

Poincarè H. (2006), *La scienza e l'ipotesi*, Milano, Bompiani.

Savage L.J. (1959), *La probabilità soggettiva nei problemi pratici della statistica in de Finetti B. (2011)*, Induzione e Statistica, Lectures given at a Summer School of the Centro Internazionale Matematico Estivo, held in Varenna (Como), Italy, June 1-10, Springer.

Vailati G. (1905), *Sull'arte di interrogare*, Rivista di Psicologia, anno I, n. 2, marzo-aprile.