



# RISK MANAGEMENT MAGAZINE

Anno 15, numero 1

Gennaio – Aprile 2020

Poste Italiane - Spedizione in abbonamento postale – 70% aut. DCB / Genova nr. 569 anno 2005

TESTATA INDIPENDENTE CHE NON PERCEPISCE CONTRIBUTI PUBBLICI (legge 250/1990)

In collaborazione con

REFINITIV

## IN QUESTO NUMERO

### ARTICOLI A CARATTERE DIVULGATIVO

<b>3</b>	Non-Performing Exposures delle banche: diktat impazienti e soluzioni nazionali vs gestione paziente e Asset Management Companies a livello europeo di Rainer Masera
<b>9</b>	NPL disposal treatment in the LGD estimates di Giacomo De Laurentis, Corrado Pavanati, Fabio Salis, Giovanna Compagnoni e Claudio Andreatta
<b>12</b>	Political risks: the “red shift” in debt sustainability analysis di Andrea Consiglio e Stavros Zenios
<b>20</b>	External fraud detection through big data: towards a pro-active operational risk management di Giacomo Petrini
<b>28</b>	Nuove policy nazionali ed internazionale: possibili implicazioni sulle Banche Italiane di Camillo Giliberto

### ARTICOLI A CARATTERE SCIENTIFICO (sottoposti a referaggio)

<b>38</b>	Classification of NPL with a Random Forest approach di Massimiliano Zanoni
<b>50</b>	Stima prospettica delle misure finanziarie di rischio mediante reti neurali dinamiche: un'applicazione al mercato statunitense di Carlo Decherchi e Pier Giuseppe Giribone

### Risk Management Magazine

Anno 15 n° 1 Gennaio - Aprile 2020

#### Direttore Responsabile:

Maurizio Vallino

#### Condirettore

Corrado Meglio

#### Consiglio scientifico

Giampaolo Gabbi (Direttore del Consiglio Scientifico), Ruggero Bertelli, Paola Bongini, Anna Bottasso, Marina Brogi, Ottavio Caligaris, Rosita Cocozza, Simona Cosma, Paola Ferretti, Andrea Giacomelli, Pier Giuseppe Giribone, Adele Grassi, Valentina Lagasio, Duccio Martelli, Laura Nieri, Pasqualina Porretta, Anna Grazia Quaranta, Francesco Saita, Enzo Scannella, Cristiana Schena, Giuseppe Torluccio.

#### Comitato di redazione

Emanuele Diquattro, Fausto Galmarini, Igor Gianfrancesco, Camillo Giliberto, Rossano Giuppa, Aldo Letizia, Enrico Moretto, Paolo Pallioli, Enzo Rocca, Fabio Salis

**Vignettista:** Silvano Gaggero

#### Proprietà, Redazione e Segreteria:

Associazione Italiana Financial Industry Risk Managers (AIFIRM), Via Sile 18, 20139 Milano

Registrazione del Tribunale di Milano n° 629 del 10/9/2004

ISSN 2612-3665

E-mail: [risk.management.magazine@aifirm.it](mailto:risk.management.magazine@aifirm.it); Tel. 389 6946315

**Stampa:** Algraphy S.n.c. - Passo Ponte Carrega 62-62r 16141 Genova

**Le opinioni espresse negli articoli impegnano unicamente la responsabilità dei rispettivi autori**

SPEDIZIONE IN ABBONAMENTO POSTALE AI SOCI AIFIRM RESIDENTI IN ITALIA, IN REGOLA CON L'ISCRIZIONE

Rivista in stampa: 8 Aprile 2020



Rivista accreditata AIDEA

## **Processo di referaggio degli articoli proposti per la pubblicazione**

Gli articoli che sono proposti alla rivista per la pubblicazione sono sottoposti in forma anonima a due successivi livelli di referaggio.

Il primo livello di referaggio (di ammissibilità) viene effettuato sull'articolo dai membri del Consiglio Scientifico che ne valutano la congruità ai temi trattati dalla rivista.

Il secondo livello di referaggio (di pubblicabilità) viene effettuato sull'articolo da due referee scelti all'interno del Consiglio Scientifico o all'esterno tra accademici, ricercatori, esperti della materia, che ne valutano il contenuto e forma.

## **Regolamento redazionale**

“Risk Management Magazine” è il periodico di AIFIRM (Associazione Italiana Financial Industry Risk Managers) ed interamente dedicato ai temi del risk management.

La sua struttura organizzativa prevede, oltre al direttore responsabile, un condirettore e un Consiglio Scientifico formato da accademici; quest'ultimo è garante della qualità e correttezza degli articoli pubblicati.

La rivista favorisce la diffusione di tutti i contenuti afferenti i temi del risk management, dagli aspetti normativi a quelli organizzativi e alle technicalities e vaglierà con interesse, per mezzo del Comitato Scientifico, i contributi che le perverranno.

Gli articoli proposti dovranno pervenire in formato Microsoft Word carattere Times New Roman 10 ed avere un numero di battute compreso tra 6.000 e 20.000; è gradita la presenza di tabelle e grafici.

La bibliografia deve essere redatta specificando accuratamente le fonti.

Si richiede la predisposizione di un Abstract in lingua inglese.

Le opinioni espresse negli articoli impegnano unicamente la responsabilità dei rispettivi autori.

Sul sito [www.aifirm.it](http://www.aifirm.it) è pubblicata la Dichiarazione sull'etica e sulle pratiche scorrette nella pubblicazione dei lavori scientifici.

# Non-Performing Exposures delle banche: diktat impazienti e soluzioni nazionali vs gestione paziente e Asset Management Companies a livello europeo

di Rainer Masera (Università Marconi)

## 1 Introduzione

In Europa e in Italia sono stati registrati significativi progressi dalle banche in materia di Non Performing Exposures (NPE: *past due, unlikely to pay* – UTP e *non performing loans* – NPL).

Nell'Euroarea gli NPL hanno toccato il picco nel 2013 (in Italia nel 2015, Sala 2019), con *unexpected losses* molto alte, ma sono in graduale discesa. Permangono elevati in alcuni paesi (Grecia, Cipro, Portogallo e Irlanda, CE 2019).

In Italia sono stati ottenuti risultati molto significativi, con riferimento sia agli aggregati lordi, sia a quelli netti (Angelini 2018, Sala 2019). È necessario proseguire lungo le direttive di riduzione. Gli NPL consumano capitale, aumentano i costi correnti e riducono la profittabilità delle banche. Il rischio di una recessione e di un nuovo dilatarsi degli NPL non è trascurabile.

Scopo di questo mio intervento non è peraltro quello di cercare di offrire un quadro di insieme e di soffermarsi sui risultati<sup>1</sup>, ma di porre in discussione taluni elementi della “saggezza convenzionale” che ha guidato i processi nell’EU. Farò riferimento a un duplice ordine di considerazioni: il confronto con l’esperienza degli Stati Uniti dopo la grande crisi finanziaria (GCF) del 2007-2009; il richiamo all’esperienza italiana della Società di gestione degli attivi (SGA) del Banco Napoli.

La BCE ha sostenuto che la riduzione significativa delle NPE in tutti i paesi dell’Euroarea è condizione necessaria per consentire il completamento della Unione Bancaria e la creazione di meccanismi comuni di garanzia/assicurazione dei depositi e di riforma delle procedure di risoluzione, oggi paleamente inadeguate, come mostrato dallo stesso Governatore della Banca d’Italia. Comunque non è accettabile la tesi che il livello degli NPL in Italia sia un problema per la stabilità finanziaria europea, come ben ha spiegato Angelini (2018).

È anche ragionevole sostenere che al double-dip del 2012-13 abbiano contribuito: la strategia di smobilizzo “impaziente” degli NPL, le perdite “forzate” sui prestiti, la stringenza del vincolo patrimoniale. Tutto ciò ha originato minor credito e spinto la spirale viziosa/deflattiva. L’enfasi microprudenziale non è stata contemplata da un’adeguata analisi macroprudenziale e dalla preoccupazione per un nuovo “tipping point”.

## 2 Le difficoltà di riassorbimento degli NPL in Europa

Hanno pesato e pesano in primo luogo le diverse performance di crescita delle economie nazionali. Si manifesta pertanto un chiaro elemento di endogeneità, troppo spesso trascurato.

Occorre inoltre riconoscere che molte banche - in particolare quelle medio piccole – hanno stentato nel creare gli skill e le procedure idonei per gestire nel modo migliore la complessa problematica, in particolare quando le NPE superavano il 20% degli impegni lordi. Progressi significativi sono comunque stati realizzati, ad esempio con il *multioriginator* per le cartolarizzazioni. Sono necessari modelli idonei di collaborazione.

Bisogna sottolineare anche la rilevanza degli impedimenti strutturali che risultano ben diversi nei paesi dell’EU. I regimi nazionali di gestione delle insolvenze, i meccanismi di protezione rispetto ad azioni di fallimento, la lunghezza delle pratiche giudiziali e delle decisioni/azioni dei tribunali hanno rappresentato oneri significativi molti diversificati (Rossano 2019).

È importante rilevare che le riforme realizzate nel biennio 2015-16 per semplificare le esecuzioni immobiliari e velocizzare le procedure di vendita degli immobili pignorati hanno avuto effetti significativi. Secondo un recente studio (Giacomelli et al. 2019), il tempo mediano necessario per la vendita degli immobili oggetto di procedure esecutive si è ridotto di quasi il 40% (da 39,5 a 23,5 mesi). È anche diminuita la durata mediana del complesso delle operazioni che precedono la vendita.

Un ulteriore fondamentale fattore viene spesso trascurato: i vincoli europei agli interventi nei singoli paesi volti a creare le cosiddette NPL - Asset Management Companies nazionali, in grado di assistere i processi, fornire garanzie o eventualmente ricapitalizzare direttamente le banche “significative” con accertate prospettive di risanamento.

L’esperienza del Fondo Atlante è rappresentativa delle difficoltà appena ricordate; l’operatività degli schemi di Garanzia Cartolarizzazione Sofferenze (GACS) seppur positiva risente dell’impostazione restrittiva della DGComp e delle prospettive comunque nazionali. Occorrerebbe viceversa disporre di *asset management companies* (AMC) pan-Euroarea in grado di operare con chiare regole del gioco, definite sulla base di considerazioni macro e micro prudenziali, accompagnate da idonee forme di garanzie europee. Come mostrerò questo è l’approccio seguito negli Stati Uniti dopo la GCF, anche come reazione agli errori precedentemente compiuti alla radice della stessa crisi, connessi alla fiducia acritica sulla efficienza/razionalità dei mercati, sulla loro capacità di autoregolarsi, sulla stabilità dei modelli stocastici di misurazione del rischio<sup>2</sup>.

Era – ed è ancora – necessario riguardare il problema in un’ottica di sistema complesso che contemperi gli schemi micro con quelli macro, che tenga conto delle diseconomie esterne delle *fire sales*, che faccia propri i modelli di “fallacia di composizione”. Questo approccio più articolato è idoneo a recepire le esigenze di una regolazione bancaria proporzionale, di fatto disattese nell’Euroarea, con costi sempre più rilevanti per le banche medio-piccole con business model tradizionale che soffrono anche delle crescenti distorsioni derivanti dalla politica monetaria di tassi di interesse negativi.

<sup>1</sup> Per un’analisi complessiva della problematica si fa riferimento a Angelini 2018, Sala 2019 e Capriglione 2019.

<sup>2</sup> Emblematici di questi errori di analisi e di policy sono gli schemi elaborati dal Premio Nobel Edward Prescott (2006) e dal Presidente della Fed Alan Greenspan (2008).

### 3 Il paradosso del de-leveraging

Una leva eccessiva – definita alla Minsky<sup>3</sup> – ovvero:

$$\text{DLR} = \text{D/A} \quad \text{dove} \quad \text{DLR} = \text{debt leverage ratio}, \text{D} = \text{debt}, \text{A} = \text{total assets}$$

è pericolosa per il singolo agente economico e per l'economia nel suo complesso. È fonte di azzardo morale e implica rischio sia idiosincratico, sia sistematico, come è avvenuto all'inizio della GCF. Ma una significativa e sostenibile riduzione della leva richiede attenta considerazione dei problemi di transizione e di aggregazione. Se tutti gli agenti, i settori e in particolare tutte le banche - anche come conseguenza delle regole di applicazione degli Standard di Basilea adottate nell'Euroarea - cercano di ridurre la leva, il risultato potrebbe essere addirittura opposto a quello desiderato come conseguenza delle *fire sales* e della riduzione della ricchezza netta degli agenti economici. Il problema può essere acuito per il settore bancario proprio in dipendenza della vendita forzata dei prestiti problematici. Il paradosso del *deleveraging*, comunemente attribuito a Minsky, è la controparte finanziaria del paradosso del risparmio.

Janet Yellen (2009), allora Presidente della Riserva Federale di San Francisco, aveva ampiamente analizzato la questione e le implicazioni di politica della regolazione nell'immediato dopo crisi. In particolare aveva sottolineato la fallacia di composizione implicita nell'assumere che istituzioni bancarie sicure, a livello di singola impresa, automaticamente implicano un sistema sicuro. Per questo la supervisione macroprudenziale può risultare diversa e deve comunque essere sovraordinata a quella microprudenziale. *Deleveraging* e *fire sales* possono creare loop perversi di retroazione con perdite generalizzate *mark-to-market*, *margin calls*, *repo runs* e *collateral shortages*.

Gli effetti risultano esaltati dalle profonde interconnessioni fra banks e shadow banks (Moe 2015). Le *fire sales* si possono manifestare sia sul mercato dei crediti bancari sia su titoli cartolarizzati ingegnerizzati da shadow banks. La loro rilevanza è stata riconosciuta negli Stati Uniti con il Dodd Frank Act (2010), ma solo parzialmente accettata in Europa, nonostante le precise indicazioni contenute nel rapporto de Larosière (2009)<sup>4</sup>. In realtà il primo a identificare e a chiarire la natura del potenziale conflitto micro/macro era stato proprio il Direttore Generale della BRI, ben prima della GCF (Crockett 2000). Al di là e al di qua dell'Atlantico permane un'insufficiente attenzione ai rischi e alla regolazione dello Shadow Banking (Moe 2015, Abad et al. 2017). In Italia viceversa il sistema di regolazione è "bank equivalent" (Gola et al. 2017).

### 4 Le azioni di policy negli US dopo la GCF per evitare le fire sales

Sulla base degli schemi di analisi sopraindicati e di un approccio olistico, l'azione del Congresso, del Tesoro, della FED e della FDIC negli Stati Uniti è stata rivolta a favorire la ricapitalizzazione delle banche, a ridurre la leva senza peraltro cadere nella trappola delle *fire sales*, ovvero di realizzare "al mercato" prezzi "di saldo", aumentando la prociclicità e paradossalmente la propagazione dei rischi e in particolare del rischio endogeno (Stein 2013).

Sintetizzo di seguito i capisaldi di questi interventi che hanno favorito la più rapida e sostenibile ripresa negli US rispetto all'Europa (per un esame più dettagliato mi permetto di fare riferimento a Masera 2011 e 2018).

La logica comune alle misure adottate è stata efficacemente sintetizzata da Martin Gruenberg (2018) come segue: "*la perdita di fiducia dei mercati ha suggerito alla Riserva Federale, al FDIC e al Tesoro degli Stati Uniti di implementare programmi di liquidità e a garantire passività delle banche che hanno mitigato l'esigenza per queste istituzioni di vendere attività a prezzi fire sales*". Sulla base di queste considerazioni il primo pilastro dei sostegni posti in essere consiste nel fatto che "*l'FDIC da sola ha fornito garanzie per \$870 miliardi*". Il secondo caposaldo degli interventi è stato rappresentato dal cosiddetto TARP (Troubled Asset Relief Program). Il Programma è stato posto in essere nell'ottobre 2008 nell'ambito dell'Emergency Economic Stabilization Act per consentire al Tesoro di favorire la stabilità dei mercati finanziari attraverso l'acquisto e le garanzie dei "Troubled Assets".

Il Congressional Budget Office è stato incaricato di predisporre Rapporti annuali sui costi per le finanze pubbliche degli interventi. Il Rapporto dell'aprile di quest'anno (CBO/TARP 2019) è, di fatto, quello conclusivo: il Programma è sostanzialmente concluso e le operazioni man mano realizzate sono giunte a termine. Si può pertanto fare un consuntivo.

A fronte degli esborsi molto rilevanti il costo dei sussidi è ridotto, in particolare rispetto ai risultati significativi ottenuti nella stabilizzazione dei mercati e degli intermediari finanziari. Come mostra la Tav. 1, il sostegno complessivo alle istituzioni finanziarie è stato pari a \$313 miliardi, dei quali 295 sono stati rimborsati. I risultati conseguiti testimoniano anche l'elevata professionalità e competenza con le quali il Programma è stato realizzato. Il terzo e più importante sostegno è intervenuto dopo il salvataggio delle due *Government Sponsored Entities* (GSE), poste sotto il controllo diretto del Governo operanti nel settore dei mutui: Fannie Mae e Freddie Mac. Fannie e Freddie – secondo la comune denominazione - hanno contribuito a svolgere un ruolo di supporto e di assicurazione dei mutui e dei rinnovati e opportunamente regolati schemi di cartolarizzazione degli stessi, sia attraverso acquisti diretti dei prestiti, sia fornendo garanzie con premi di mercato. Le *Mortgage Backed Securities* (MBS) sono poi state vendute a operatori finanziari specializzati. L'intero settore è stato posto sotto il controllo della Federal Housing Finance Agency (FHFA), creata nel 2008 con lo Housing and Economic Recovery Act. La FHFA è anche responsabile del Federal Home Loan System.

<sup>3</sup> Altre definizioni comunemente utilizzate sono:

$\text{LR} = \text{C/A} = \text{C/RWA} \times \text{RWA/A} = \text{RWR} \times \text{DR}$  dove C = capitale; RWR = risk weighted capital ratio e DR = density ratio.

LM = A/C dove LM = leverage multiple.

<sup>4</sup> Sui condizionamenti europei alle strategie di gestione dei crediti deteriorati in Italia cfr. Capriglione 2017, Troiano e Uda 2018, Rossano 2019.

Tav. 1

**Actual and Projected Cash Disbursements of the Troubled Asset Relief Program**

Billions of Dollars

	Results to Date for Principal Disbursed				Additional Disbursements Anticipated
	Principal Disbursed	Repaid	Written Off <sup>a</sup>	Outstanding	
Support for Financial Institutions					
Capital Purchase Program	205	200	5	*	0
Additional assistance to Citigroup and Bank of America <sup>b</sup>	40	40	0	0	0
Community Development Capital Initiative	1	1	*	*	0
Assistance to American International Group	68	54	13	0	0
Subtotal	313	295	19	*	0
Assistance to the Automotive Industry	80	63	17	0	0
Investment Partnerships					
Term Asset-Backed Securities Loan Facility <sup>c</sup>	*	*	0	0	0
Public-Private Investment Program	19	19	0	0	0
SBA 7(a) Purchase Program	*	*	*	0	0
Subtotal	19	19	*	0	0
Mortgage Programs <sup>d</sup>	29	0	29	0	2
<b>Total</b>	<b>441</b>	<b>377</b>	<b>65</b>	*	<b>2</b>

Sources: Congressional Budget Office; Department of the Treasury.

Transactions are as of February 28, 2019.

SBA = Small Business Administration; \* = between zero and \$500 million.

a. Includes realized losses on sales of common stock. For mortgage programs, the amount represents disbursements to loan servicers.

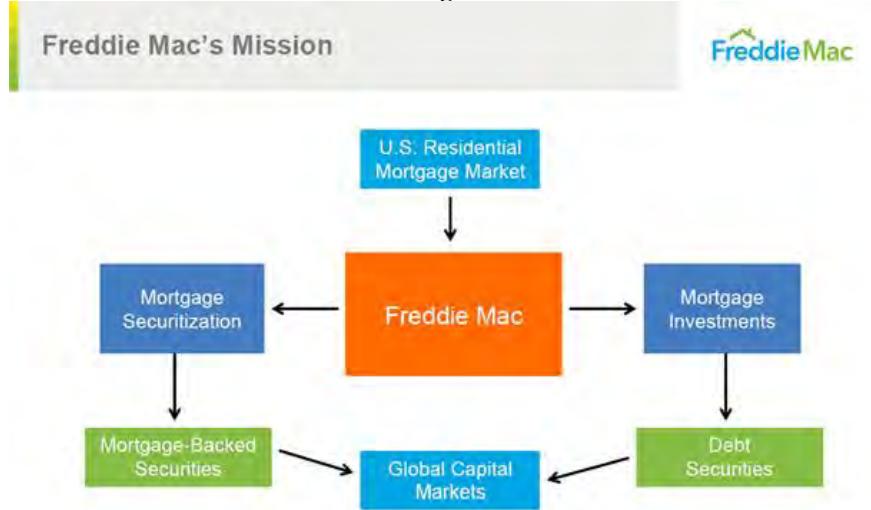
b. The Treasury also agreed to provide \$5 billion to cover potential losses on Citigroup's assets; however, those losses did not occur, so no disbursements were made.

c. The Treasury committed \$4 billion to absorb losses on loans made by the Federal Reserve through the Term Asset-Backed Securities Loan Facility; however, no losses occurred and the Treasury recouped all of the \$100 million in initial funding.

d. Of the \$50 billion initially announced for the mortgage modification programs, which include funding for state housing finance agencies and the Federal Housing Administration, \$31 billion eventually will be disbursed, CBO estimates.

Il totale dei mutui gestiti da Fannie e Freddie supera i \$6 trillion, con evidenti sostegni diretti e indiretti alle banche. Le Figg. 1, 2, 3 e 4 di seguito riportate pongono in evidenza sinteticamente i processi descritti. In particolare, come si vede dalla Fig. 3, nell'ambito del QE la Fed ha acquisito ammontari molto rilevanti di MBS, contribuendo direttamente a dare respiro al sistema bancario e finanziario ed evitando svendite forzate di attività bancarie. Jeremy Stein (2013), fino a poco tempo fa Governatore della Fed, ha mostrato che le *fire sales* forzate di attività bancarie rappresentano una modalità molto rilevante di trasmissione e amplificazione dei rischi e non solo di *derisking*.

Fig. 1



"A primary purpose is to provide stability in the secondary market for home mortgages including mortgages securing housing for low and moderate income families. This can be accomplished through both portfolio purchasing and selling activities, as well as through the securitization of home mortgages."

Source: Freddie Mac (2019)

**Fig. 2**



Source: Freddie Mac (2019)

**Fig. 3**

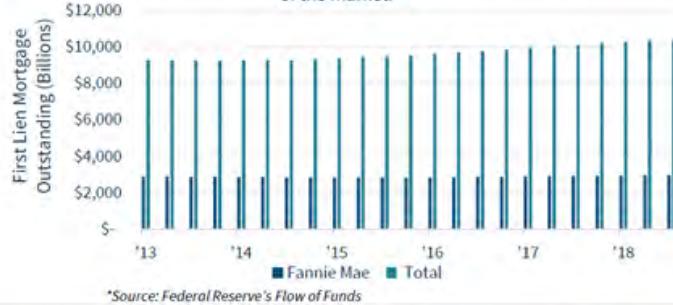


Source: Freddie Mac, Fannie Mae, Federal Reserve, Inside MBS & ABS, National Credit Union Administration and the U.S. Treasury Department.  
As of March 31, 2019.

Source: Freddie Mac (2019)

**Fig. 4**

As of March 2019, U.S. Single Family 1<sup>st</sup> Lien mortgage debt outstanding totaled \$10.4 trillion. Fannie Mae's share stood at approximately \$2.9 trillion, nearly 28% of the market.



Fannie Mae was the largest issuer of single-family mortgage securities in the first quarter of 2019.

Source: Fannie Mae (2019)

Per completare il quadro dei sostegni forniti dal Governo degli Stati Uniti ai prestiti delle banche dopo la CGF è opportuno ricordare anche l'operatività della Small Business Administration (SBA). Questa Agenzia federale offre supporto alle piccole-medie imprese (PMI), principalmente fornendo in partnership con banche (segnatamente Community Banks) e Credit Unions prestiti alle PMI assistiti da garanzie del Governo. Subito dopo la crisi con il Recovery Act e il Small Business Jobs Act la percentuale massima di garanzia è stata portata nel 2008 al 90% dei singoli nuovi prestiti (per un esame dell'operatività della SBA cfr. USA gov SBA 2019).

## 5 L'esperienza della Società Gestione di Attività del Banco Napoli

Mi avvio a concludere queste note tornando in Italia, con un flashback sull'esperienza, anche diretta, SGA/Banco Napoli. Per consentire la "privatizzazione" e la vendita del Banco, si creò nel 1997 una *bad bank* che comprendeva tutte le esposizioni non-performing (analogamente alla definizione da cui ho preso le mosse).

Sottolineo il nome: Società per la gestione di attività (oggi cambiato in AMCO – *Asset management Company*). La società viene affidata a due eccellenti banchieri della scuola del Credito Romagnolo, Marcello Valignani e Roberto Romagnoli. Viene garantita la piena indipendenza del loro operare e dello staff. È assicurato l'orizzonte temporale idoneo per gestire circa 30mila pratiche in sofferenza caratterizzate da situazioni molto complesse. La SGA disponeva comunque del cordone di sicurezza previsto dal cosiddetto Decreto Sindona, che consentiva crediti da Banca d'Italia a tassi agevolati rispetto al mercato. Il divario consentiva la ricapitalizzazione della stessa SGA.

Il Decreto Sindona del 27 settembre 1974 fu utilizzato sia per la liquidazione di Banca Unione e Banca Privata Finanziaria, facenti capo a Sindona, sia successivamente per il Banco Ambrosiano di Roberto Calvi nel 1982. Nel 2002, quando San Paolo IMI (SPIMI) ha acquisito Banco Napoli, l'operazione è stata studiata e realizzata dal Dottor Arcuti e da chi vi parla solo dopo un'attenta analisi e "due diligence" della società, che poi confluiscce in SPIMI.

Il bilancio di SGA è riportato nella Tav. 2 qui allegata che sottolinea i successi ottenuti.

Tav. 2 – Ultimo bilancio di SGA (2015)

Tipologia di attività	Prezzo pagato alla data del 31-12-1996	RECUPERI									TOTALE AL 31/12/2015
		DAL 1997 AL 2000	DAL 2001 AL 2005	DAL 2006 AL 2010	2011	2012	2013	2014	2015	(A+B+C+D+E+F+G+H)	
	A)	B)	C)	D)	E)	F)	G)	H)			
Crediti	6.272,5	2.222,5	1.533,8	943,4	83,5	69,4	76,1	61,9	44,9	5.035,7	
Titoli e partecipazioni Italia (compreso BNI)	153,0	162,5	56,9	24,2	2,0	1,5	-	1,0	-	248,2	
Transazioni con ex Banco di Napoli (*)	-	12,9	125,3	13,9	-	-	-	-	-	152,1	
<b>Totale attivi ex Banco di Napoli</b>	<b>6.425,9</b>	<b>2.398,0</b>	<b>1.716,0</b>	<b>981,5</b>	<b>85,5</b>	<b>70,9</b>	<b>76,3</b>	<b>62,9</b>	<b>44,9</b>	<b>5.436,0</b>	

Fonti: Verdelli (2017), Repubblica Economia&Finanza / Camera di Commercio di Napoli

Come è noto si è aperto un dibattito sulla rilevanza del sostegno finanziario di Banca d'Italia per spiegare gli eccellenti risultati (Servizio bilancio della Camera dei Deputati 2016, Zingales 2016, Giannola 2016). Non voglio entrare in questa polemica, ma sottolineo la rapidità e l'efficienza di Banca d'Italia con particolare riferimento alla gestione del Caso Banco Ambrosiano, che ho avuto modo di seguire sotto la guida lungimirante ed estremamente efficace del Governatore Carlo Azeglio Ciampi (Masera 2019a).

## 6 Conclusioni

Significativi progressi sono stati ottenuti in Europa nel necessario processo di riassorbimento delle NPE e segnatamente degli NPL. La questione degli UPT resta sotto certi versi irrisolta (Capriglione 2019). Peraltro, come indicato, i costi dell'approccio seguito fatti ricadere sulle banche e sull'economia in termini di flussi di credito sono stati notevoli, anche per la declinazione adottata degli standard di Basilea (Szego 2004, Roncaglia 2012, Masera 2019b).

Il confronto sia con gli Stati Uniti dopo la GCF, sia con l'esperienza italiana del passato, consente di affermare che gli schemi adottati nell'UE sulla gestione dei crediti deteriorati sono stati caratterizzati da un trade-off insoddisfacente fra benefici e costi (cfr. anche de Larosière, 2013). Se si dovesse verificare una nuova recessione, come ritengo probabile, le considerazioni qui svolte avrebbero valenza non solo come diagnosi rivolta al passato. Comunque maggior attenzione va posta sulla proporzionalità nella regolazione non solo tra banche grandi e medio-piccole, ma anche tra banche e shadow banks con particolare riferimento ai *leveraged loans*.

Rainer Masera

## Bibliografia

- Abad J. et al. (2017). Mapping the interconnectedness between EU banks and shadow banking entities, ESRB, March.
- Angelini P. (2018). I crediti deteriorati: mercato, regole e rafforzamento del sistema, Intervento del Vice Capo del Dipartimento Vigilanza bancaria e finanziaria della Banca d'Italia, Roma, 12 ottobre.
- Capriglione F. (2017). La nuova gestione delle crisi bancarie tra complessità normativa e logiche di mercato, Rivista trimestrale di diritto e economia, 1.
- \_\_\_\_ (2019). La problematica dei crediti deteriorati, Relazione di base, Convegno sugli NPL, Università Parthenope, Napoli, 17-18 ottobre.
- Commissione europea (2019). Comunicazione della Commissione Comunicazione della Commissione al Parlamento Europeo, al Consiglio Europeo, al Consiglio e alla Banca Centrale Europea, Quarta relazione sui progressi compiuti nella riduzione dei crediti deteriorati e nell'ulteriore riduzione del rischio nell'Unione bancaria, Brussels, 12 giugno.
- Congressional Budget Office (2019). Report on the Troubled Asset Relief Program, April.
- Crockett A. (2000). Marrying the micro-and macroprudential dimensions of financial stability, BIS Speeches, September 21.
- de Larosière, J. (2013). The trade-off between bank regulation and economic growth, Central Banking Journal, 28 February.
- de Larosière J., Balcerowicz L., Issing O., Masera R., McCarthy C., Nyberg L., Pérez J., Ruding O. (2009). The High-level Group on Financial Supervision in the EU, Report. Brussels, February 25.
- Dodd-Frank Wall Street Reform and Consumer Protection Act (2010). (4173) Washington (DC): HR.
- Fannie Mae (2019). Investor Presentation, August.
- Federal Housing Finance Agency (2019). FHFA at a glance.
- \_\_\_\_ (2019b). Fannie Mae and Freddie Mac Oversight, Washington D.C., June.
- Freddie Mac (2019). Investor Presentation, August.
- Giacomelli S. et al. (2019). Le procedure esecutive immobiliari: gli effetti delle riforme del 2015-16 sulla durata delle procedure, Note di stabilità finanziaria e vigilanza n. 16, Banca d'Italia, Roma, ottobre.
- Giannola A. (2016). Banco di Napoli, un'eredità contesa, Il Sole24Ore, 27 agosto.
- Gola C. et al. (2017). Shadow banking fuori dall'ombra: l'intermediazione non bancaria e il quadro regolamentare italiano, Banca d'Italia, wp 372, febbraio.
- Greenspan A. (2008). I made a mistake, Hearings, Capitol Hill, Washington DC, October 23.
- Gruenberg M.J. (2018). An Essential Post-Crisis Reform Should Not Be Weakened: The Enhanced Supplementary Leverage Capital Ratio, Peterson Institute for International Economics, Washington, September 6.
- KPMG (2018). Non-performing loans in Europe. What are the solutions?, available at <https://assets.kpmg/content/dam/kpmg/xx/pdf/2017/05/non-performing-loans-in-europe.pdf>, August.
- Marchesano M. (2016). Il caso Sga, GoWare.
- Masera, R. (2012). Risk, Regulation and Supervision of Financial Systems: US and Eurozone Solutions. ZÖR, 67: 251–280.
- \_\_\_\_ (2018a). Political economy of liquidity: the European Economic and Monetary Union, Palgrave, London.
- \_\_\_\_ (2018b). Bu Rules on Banking Crises: The Need for Revision, Simplification and Completion, in Troiano V. and Uda G. (2018).
- \_\_\_\_ (2019a). Il Governatore Ciampi e la soluzione del “Caso Ambrosiano”: profili internazionali di vigilanza e di politica monetaria, Carlo Azeglio Ciampi Governatore della Banca d'Italia, Banca d'Italia, Roma, 9 luglio.
- \_\_\_\_ (2019b). Leverage and risk-weighted capital in banking regulation, Aifirm, Roma, 21 giugno.
- Masera R. and Mazzoni, G. (2016). On the non-neutrality of the financing policy and the capital regulation of banking firms. Studies in Economics and Finance, Vol. 33 Iss 4 pp. 466 - 487: <http://dx.doi.org/10.1108/SEF-09-2014-0179>
- Moe T. (2015). Shadow banking: policy challenges for central banks, Norges Bank, Journal of Financial Perspectives, Vol. 3, Issue 2, July.
- Prescott E. (2006). The Golden Age of Economics, Trinity University in San Antonio, April.
- Repubblica (2017). Economia&Finanza, 28 giugno.
- Roncaglia A. (2012). Keynesian uncertainty and the shaky foundations of statistical risk assessment models, PSL Quarterly Review, vol. 65 n. 263, pp. 437-54, December.
- Rossano D. (2019). La gestione dei crediti deteriorati nelle crisi bancarie, Convegno sugli NPL, Università Parthenope, Napoli, 17-18 ottobre.
- Sala G. (2019). La gestione degli NPL. Nuove prospettive regolamentari e di mercato, Aifirm, Palermo, 4 ottobre.
- SBA (2019). <https://www.sba.gov/>.
- Servizio bilancio della Camera dei Deputati (2016). Bilancio Sga, giugno.
- Stein J. (2013). The Fire Sales Problem and Securities Financing Transactions, Federal Reserve Bank of Chicago, November 7.
- Szego G. ed. (2004). Risk measures for the 21<sup>th</sup> Century, Wiley, London.
- Troiano V. e Uda G. (2018). La gestione delle crisi bancarie, Wolters Kluwer, Milano.
- Yellen J. (2009). A Minsky Meltdown: Lessons for Central Bankers, FRBSF Economic Letter, May 15.
- Verdelli C. (2017). Il ritorno della Sga: la Bad bank del Banco di Napoli e le speranze del Tesoro, La Repubblica, 28 giugno.
- McWilliams J. (2019). Remarks, Federal Deposit Insurance Corporation (FDIC), Washington (DC), June 4.
- Zingales L. (2016). Quel «tesoretto» della bad bank del Banco di Napoli, , Il Sole24Ore, 20 agosto.

# NPL disposal treatment in the LGD estimates

di Giacomo De Laurentis (Bocconi University), Corrado Pavanati (UniCredit), Fabio Salis (Creval), Giovanna Compagnoni (Mediobanca), Claudio Andreatta (UBI Banca)

## Abstract

The financial crisis resulted in an exponential increase of Non-Performing Loans (NPL) within banks' portfolios, which hugely impacted their profitability.

To push banks to get rid of the legacy of the crisis, regulators issued a number of regulations bearing significant implications for banks.

Among other things, a threshold with respect to the ratio of Non-Performing Exposure to the bank's overall Exposure (i.e. NPE ratio) has been set, past which banks become subject to a specific monitoring regime.

The regulatory stance coupled with financial market pressure to lead banks to sell off material portions of their Non-Performing Loans portfolios.

Since 2017 Italian banks started disposing NPL portfolios in a growing number. Disposals initially concerned portfolios of bad loans (i.e. *sofferenze*) to then embrace also Unlikely-to-Pay loans.

For AIRB banks (i.e. banks allowed to use internal models for capital requirements purposes) the inclusion of portfolio/massive NPL disposals into their Loss Given Default (LGD) estimates would raise capital requirements to unprecedented levels.

More in detail, once massive NPL disposals are included into LGD models the expected losses estimates jump to levels that are representative of forced sell-off experiences and thus become detached from the internal workout processes used for newly defaulted loans. If used for pricing, provisioning and capital quantification purposes, such biased estimates would impair the formation of new credit, with negative macroeconomic consequences.

To handle the impact of the NPL massive disposals on their LGD models, AIRB banks are working on a number of methodological approaches, that either factor in the probability that massive disposal will occur in the future or try to mitigate the extraordinary disposal prices. Having acknowledged the macroeconomic impact of the inclusion of NPL disposals into LGD models, the European Parliament provided AIRB banks with a limited leeway in their modelling practices.

A specific article of the new Capital Requirement Regulation deals with this issue: article 500. This regulatory evolution is aimed at offsetting the distortive impact that the massive disposals, characterized by the disposal price, would have on the LGD parameter estimation. In fact the article allows to assume that positions object of massive disposal follow the same recovery pattern observed for workouts, as it happens for "open positions", thus offsetting the actual anomalous recovery pattern that would introduce bias in the estimate. Different aspects of the methodology related to the implementation of article 500 continue to be examined by AIRB Italian banks.

## 1 Economic context

With the advent of Financial and Economic Crisis in the Eurozone in 2008 came a rise in the financial sector NPL portfolios, impacting all types of portfolios. The effect of the economic crisis was especially potent on a number of EU bodies, notably Italy, whose banking sector saw a predominantly linear increase of its NPL stock for almost a decade.

In 2015 the gross number of NPL, excluding bad loans, rose to € 350 billion, causing the NPL ratio (as a proportion of Total Loans) to reach historical levels (approx. 17%)<sup>1</sup>. This large quantity of NPL has a number of negative side effects on financial institutions, including a decrease in asset quality and limitations on access to capital, which in turn hinders lending capacity and overall profitability.

Due to the potentially harmful ramifications on both financial institutions and consumers, regulatory bodies, notably the European Central Bank (ECB) and European Banking Authority (EBA) were forced to intervene and implement various measures to ensure the proper management and treatment of these non-performing exposures.

An integral part of the guidance and expectations set by the regulators included a suggested quantitative threshold or trigger (5% of gross loans), which incentivized financial institutions to effectively run off their NPL portfolios on a sizable and expedited basis.

However, illiquid characteristics of the NPL market resulted in financial institutions observing material differences between the fair values and selling prices of positions belonging to the NPL portfolios, impacting both bank financial statements and capital ratios with a direct link to overall business sustainability.

The computation of the regulatory capital is based on the estimates of the LGD models, and their respective outputs. Ultimately, the risk differentiation and quantification of these models are affected by the distortive nature of the NPL portfolio. A number of methodological solutions have been formulated in order to mitigate or offset any discrepancies created within the model by the introduction of disposed positions in the portfolio used as sample for LGD estimation.

While a new regime of proactive regulatory measure seeks to create a more structured approach surrounding the management of impaired loans, including implementation of proper governance, risk management, monitoring, and recovery planning, the LGD parameters remain a crucial element impacted by the effect of NPL portfolio disposals.

<sup>1</sup> Bank of Italy – Financial Stability Report No. 1/2019

## 2 Probability of disposal

Ultimately, the LGD is computed as an average of the loss rate associated to the disposal event and the loss rate associated to the workout scenario, weighted by the respective probability of occurrence.

In doing so the LGD computation follows the prescriptions of the recent EBA Guidelines on the PD estimation, LGD estimation and treatment of defaulted assets that requires banks to use LGD estimation methods that take into account various recovery scenarios and their probability of occurrence.

The *Probability of disposal* is estimated through a *hybrid* approach, as it makes use of both historical NPL data of the bank, as well as forward-looking simulations of potential future macroeconomic conditions, subject to two primary assumptions:

- General economic conditions affect banks' NPL volumes (i.e. worse economic conditions lead to higher levels of NPL)
- when NPL exceeds certain thresholds banks are bound to sell a portion of the NPLs through extraordinary disposals (i.e. typically priced at a discount due to market illiquidity)

The forward-looking approach aims at simulating future macroeconomic conditions leveraging models and experience banks have developed when implementing the IFRS9 accounting principle. These simulations allow to incorporate potential portfolio disposal, either already planned or not yet foreseen, into LGD estimates.

The *Probability of disposal* estimation utilizes an end-to-end approach, consisting of the following steps:

- 1) Identification of the relationship between one or more macroeconomic variables (e.g. GDP, interest rates, etc.) and the NPL ratio through an econometric model.

Such relation can be investigated by means of different econometric models, such as linear regression or Vector Autoregressive Model (VAR).

With reference to the independent variables, the long list is identified leveraging the already existing models that make use of macroeconomic variables, first and foremost those necessary for implementing the IFRS 9 principle. With reference to the dependent variable, different forms of the NPL stock/ratio can be considered: absolute and relative terms (i.e. differences or variations).

The long list of independent variables is analyzed through both univariate and correlation analyses, in order to move to the identification of the economic indicator(s) to be included in the final econometric model.

- 2) Simulation (i.e. Monte Carlo) of future macroeconomic indicators.

The starting point of the simulation is identified by applying the estimated parameters of the macroeconomic model to the current NPL level (in absolute or relative terms).

Both the number of simulations and the time horizon are elements that have to be identified prior to the beginning of the simulation.<sup>2</sup>

At each step of the simulation, an innovation/error is generated in order to project a step forward the simulated variable (by applying the parameters estimated to the selected macroeconomic variable).

At the end of each simulated year, if the simulated NPL level is higher than the time varying threshold, a disposal is assumed to occur and the NPL level is set back to the value of the threshold.

Both the step of the projection of the macroeconomic variable one step forward and the check of the NPL level are repeated until the end of the time horizon.

A path of simulated NPL levels is computed for each simulation.

The process is repeated for each simulation.

- 3) Identification of the threshold above which the bank is forced to dispose part of its NPL portfolio.

This step is required in order to identify the need of selling part of the NPL portfolio, define the amount of NPL to be disposed, set the target NPL level to be reached within a pre-defined time horizon.

- 4) Computation, for each period of time (year) of two components: (1) the probability that the NPL value breaches the threshold and (2) the portion of portfolio to be disposed in order to bring back the NPL value under the threshold.

- 5) Calculation of the probability of disposal over the entire identified time horizon.

For each year, the probability of disposal is the product of the probability of a scenario to overcome the time-varying threshold and the average amount to be sold to bring back the NPL level to the threshold.

The product of the yearly probabilities of disposal is the overall probability of disposal.

<sup>2</sup> The time horizon can be identified in line with the characteristics of both the application portfolio and the recovery process. Also, the industrial plan can be taken into account.

- 6) Estimate of a “stressed” probability of disposal or of a “downturn value” of the probability of disposal, i.e. using downturn values of the future macroeconomic conditions.

One possible approach is to compute the stressed value of the probability of disposal by considering a negative scenario of the macroeconomic variables and to include it in the simulation, in order to get a “downturn value” of the probability of disposal.

Given potential shortcomings of model performance, including the lack of a sufficiently long time series, it's also considered a best practice to have a back-testing framework in place.

### **3 Mitigation coefficient**

An alternative approach to the “Probability of Disposal” approach, may consist of applying a “Mitigation Coefficient”, or specific weight to transactions sold via extraordinary NPL disposal. This approach allows banks to include all default experiences in the sample used as basis for the LGD estimation while limiting the potential bias from the NPL disposal.

The approach relies on the quantification of a so-called “Mitigation Coefficient” to be combined with the workout LGD in the model estimation phase, for instance through the implementation of a weighted regression (Weighted OLS) or in the LGD look-up table. When the approach is aimed at offsetting up to the maximum extent possible the impact of one or more extraordinary NPL disposal, the Mitigation Coefficient is computed as an Adjustment Factor, which will be based on the distance between the economic loss observed on extraordinary disposals and the one observed on similar transactions managed through the internal workout process or ordinary disposals (where in ordinary conditions NPL disposals can be considered a natural alternative to internal workout if as efficient as the last one). Once computed, the Factor applies to the concerned transactions as long as they are included in the LGD estimation sample; i.e. new NPL disposals would not change the previously computed Adjustment, that will be freezed. Differently, when the mitigation procedure is aimed at capturing both the extraordinary nature of NPL disposals and their probability of occurrence, two dimensions must be considered: the Adjustment Factor (described above) and the Baseline Mitigation Coefficient (calculated as the ratio between the probability of disposal in an ordinary environment, i.e. low NPL levels, and the probability of disposal in extraordinary conditions, i.e. high NPL levels).

### **4 Interpretation of targeted amendments to credit risk requirements on portfolio disposals designed to limit the impact on capital requirements**

European Commission recently published a new Regulation, the so-called CRR2, which contains a specific article (i.e. Article 500) that takes into account the possibility to mitigate the effect that the massive disposal of NPL has on the LGD models and their respective outputs.

These new guidelines continue to be examined and interpreted by industry experts, though, the article 500 opens a door for banks to avoid taking the brunt of the LGD impact due to the NPL sell-offs.

The article basically allows institutions to “[...] *adjust their LGD estimates by partly or fully offsetting the effect of massive disposals of defaulted exposures on realised LGDs up to the difference between the average estimated LGDs for comparable exposures in default that have not been finally liquidated and the average realized LGDs including on the basis of the losses realized due to massive disposals [...]*”. Although the principle stated by the European Parliament is clear, the rule text leaves room for different interpretations about the methods banks may use to partly or fully offset the effect of massive NPL disposals on LGD estimations.

In this regard, two methodological approaches have been singled out, both suitable to implement the article 500 in the LGD estimation framework. The approaches share similar background rationale, i.e., that disposed positions should be treated as open ones, though each are interpreted slightly differently. In order for both approaches to be applied, institutions are required to adopt a methodology which includes open positions (so called incomplete recovery processes) in the respective LGD model – in order to comply with EBA Guidelines. In particular, under Approach 1 the disposal adjustment is estimated in line with the incomplete recovery processes treatment, basically meaning that disposals are treated as if they were not been disposed, but were still open, with the relative projection of future recoveries. Approach 2 is based on the interpretation of the “not finally liquidated” concept as referred to positions still open in the balance sheet of the Institution, leading to an adjustment based on the application of the estimated LGD values to those positions. Nonetheless, both approaches lead to an LGD calibrated according to a corrective factor for massive disposals to be estimated (i.e. includes closed, substantially closed, open positions, and massive disposals). It is clarified that the adjustment applied on disposed exposures eligible for article 500 treatment should be frozen at the moment of first quantification.

Giacomo De Laurentis, Corrado Pavanati, Fabio Salis, Giovanna Compagnoni, Claudio Andreatta

# Political risks: the “red shift” in debt sustainability analysis

di Andrea Consiglio and Stavros Zenios<sup>1</sup>

(This is an expanded version of a blog posted at Bruegel on January 22, 2020, <https://bruegel.org/2020/01/incorporating-political-risks-into-debt-sustainability-analysis/> )

## Abstract

Political stability and economic policy uncertainty can be key determinants of sovereign debt dynamics, and we show how they can be incorporated in debt sustainability analysis.

We distinguish between short-term ambiguity and long-term uncertainty about political risk factors, and using a combination of narrative scenarios and calibrated probabilistic scenarios we obtain a comprehensive heatmap of high-risk debt dynamics. We use Italy as an interesting case study and demonstrate a “red shift” in the assessment of vulnerabilities when accounting for political risks. *Ignoring these risks can lead to excessive optimism and wrong decisions.*

-----  
La stabilità del sistema politico istituzionale e l'incertezza riguardo le politiche economiche sono due fattori chiave che possono influenzare la dinamica del debito pubblico.

Nell'articolo si propone un modello di analisi della sostenibilità del debito sovrano che tenga conto dei fattori di rischio concernenti l'assetto istituzionale di un paese e le sue politiche economiche.

In particolare, distinguendo fra ambiguità a breve termine e incertezza a lungo termine dei fattori di rischio politico e, utilizzando una combinazione di scenari narrativi e scenari probabilistici, si costruisce una "heatmap" che permette di attribuire ad ogni politica fiscale la probabilità che l'obiettivo di riduzione dello stock di debito o del deficit sia soddisfatto.

Il modello è applicato al caso Italia. I risultati mostrano un "red shift" della vulnerabilità del debito pubblico italiano quando sono inclusi nell'analisi i fattori di rischio politico.

Si può quindi concludere che ignorare i rischi derivanti dall'instabilità del sistema politico-istituzionale, o quelli derivanti dall'incertezza delle politiche economiche, può condurre a un eccessivo ottimismo e a conseguenti scelte sbagliate.

## 1 Introduction

---

“Partly in reaction to the European experience, the traditional approach to debt sustainability assessment has evolved. A fresh view on things is one way to improve the analysis of debt sustainability and develop best practices.”

(Klaus Regling, Managing Director ESM, Debt sustainability conference, Dec. 2018.)

---

The Greek debt crisis revealed a significant concern regarding the effectiveness of traditional debt sustainability analysis (DSA). Such analysis is used to assess whether a country meets the criteria for receiving international assistance, but around crisis episodes uncertainty is high and focusing on average dynamics, or on a few scenarios, can conceal potential risks.

This is especially acute when official lending moves into addressing problems of economies with large and active bond markets with adverse feedback loops. And whereas DSA applies to crisis countries, an early warning system identifying vulnerabilities is relevant for all countries. We suggest that a more general, less stringent, debt vulnerabilities analysis (DVA) could be used to assess a country’s debt management policies to identify vulnerabilities, without leading immediately to policy consequence.

DVA would not carry the significant connotations of DSA. For instance, the Dutch State Treasury Agency evaluates its public debt management practices every three years, even as its debt-to-GDP ratio is only 50%. The Agency carries out a comprehensive analysis prior to a political review, including an evaluation of vulnerabilities, and the policy implications are transmitted by the Dutch Minister of Finance to the Parliament.<sup>2</sup> DVA would have less stringent criteria for raising a red

---

<sup>1</sup> A. Consiglio is with University of Palermo, Palermo, Italy. S.A. Zenios is with University of Cyprus and Non-Resident Fellow of Bruegel. We acknowledge useful comments from Marilena Athanasopoulou, Maria Demertzis, Rebecca Christie, and Guntram Wolff. We thank Giovanni Pagliardi for comments and for permission to use data from a draft paper on the sensitivity of bond yields to political risk factors.

<sup>2</sup> The evaluation for the 2019 political review was carried out by SEO Amsterdam Economics, and their report is available as J. Hers, R. Beetsma, J. Witteman, and N. Verheuvel. *Independent assessment of the 2016–2019 interest rate risk framework and funding policy*, Report commissioned by Dutch State Treasury Agency, SEO Amsterdam Economics, Amsterdam, February 2019. The full report (in English) and

flag, such as when a country may pass the DSA test but its debt is non-decreasing from very high levels. Also, DVA should cover a broad set of risk factors to identify problems before they become critical. Italy presents a good example where high debt levels leave the country vulnerable to market reactions, with a real risk of entering unsustainable territory if remedial actions are not taken.

DSA tools have been evolving to assess risks beyond mean value projections. The IMF adopts a “fan-chart approach” to debt sustainability (IMF, 2013), the ECB embeds debt simulations under a benchmark and several narrative scenarios (Bouabdalah et al., 2017), and ESM optimizes the sovereign’s tail risk to assess, with high probability, whether non-sustainable debt dynamics are imminent or a future possibility (Zenios et al. 2019).

In these approaches, uncertainty (i.e., the probability distribution of key risk factors) takes centre stage. IMF advocates the use of a baseline scenario with alternative narrative scenarios, and for “higher scrutiny countries” it supplements its analysis with fan charts. The ECB framework also provides for a stochastic analysis using a probabilistic approach with heat maps, in addition to narrative shock scenarios. ESM uses calibrated multi-period scenario trees to derive fan charts of both debt stock and flow dynamics, and optimizes the bond issuance to limit extreme (tail) adverse events.

## 2 Which risk factors?

The main risk factors are common to all three institutions: They include the fiscal consolidation path, GDP growth, and financial assumptions relating to the sovereign bond yields. Narrative scenarios can include ageing costs, macro (bank) stress tests, inflation shocks, structural shocks, contingent liabilities, and privatization receipts. *We argue that political risk factors can also be quantified, and should be part of debt analysis.*

Political risks guide expert judgment by the institutions, but they are not part of DSA models and are treated without the quantitative rigour reserved for macroeconomic, financial, and fiscal risks. IMF makes references to political risks and policy uncertainty in its Article IV Consultation reports.<sup>3</sup> Such references appear twenty-six times in the 2016 Article IV report for Greece, twelve times in the 2018 report, and four times in 2019. The ESM uses governance and/or political risk ratings — from the World Bank, the Corruption Perception Index of Transparency International, the aggregate political risk index of the PRS group — in its Sovereign Vulnerabilities Index and, likewise, ECB uses such ratings to generate a heat map, classifying RED countries in the bottom ratings tercile, GREEN in the top tercile, and YELLOW in the middle. Such broad treatment of political risks is useful, but unlikely to be effective.

The omission of political risks from DSA models for distressed countries is surprising, since political risks are clearly perceived as important. The lack of quantitative rigour is, in our opinion, due to inadequate quantifiable empirical evidence and an appropriate framework, but in recent years both have become available, allowing us for a fresh view on the problem. Inadequate treatment of these risks can lead to excessive optimism and wrong decisions, as we will demonstrate for the case of Italy.

The systematic quantification of political risks has been receiving increasing attention. The economic effects of politics have been studied as early as Schattschneider (1935), and the market effects of political uncertainty goes back to the study of Imperial Germany and the Weimar Republic (Bittlingmayer, 1988). Recent, renewed, interest is driven in part by the compilation of databases that facilitate cross-sectional studies. Such databases include the Ifo World Economic Survey-WES<sup>4</sup> (25 years of semi-annual data for 66 countries), the World Bank (25 years of annual data for 214 countries), the ICRG index (40 years of annual data for up to 140 countries), and the Economic Policy Uncertainty index<sup>5</sup> posted on the Dallas Fed (25 years of monthly data for 21 countries). Impetus was also given by theoretical models and empirical evidence that the markets price political risks.<sup>6</sup>

We show how political risks can be incorporated in DSA (and DVA) and materially affect the conclusions. First, we identify two key quantifiable dimensions of political risk. Second, we make a distinction between short-term ambiguity about the political factors that cannot be measured, and long-term risks that are modeled probabilistically. Third, we show how a

---

the letter of the Finance Minister (in Dutch) to the Parliament are available at <https://www.dst.nl/actueel/persberichten/2019/11/22/nieuw-beleidskader-financiering-staatsschuld>. For DVA they use the model of Zenios et al. (2019), see pp. 19, 24-30 in the SEO report.

<sup>3</sup> See, for instance, the 2018 Financial Stability Report and the 2018 Global Outlook Reports

<sup>4</sup> Becker, S. O. and K. Wohlrabe. “Micro data at the Ifo Institute for Economic Research: The Ifo Business Survey, usage and access,” Working Paper 47, Ifo Institute for Economic Research, Munich, 2007.

<sup>5</sup> Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty, *The Quarterly J. of Economics*, 131:1593–1636.

<sup>6</sup> See Pastor, L. and P. Veronesi (2012). Uncertainty about government policy and stock prices, *The Journal of Finance*, 67:1219–1264, and Bekaert, G., C. R. Harvey, C. T. Lundblad, and S. Siegel (2014): Political risk spreads, *Journal of International Business Studies*, 45:471–493.

combination of narrative scenarios about the short-term ambiguity, and calibrated probabilistic scenarios for long-term risks, gives a comprehensive heatmap of *high-risk debt dynamics*.

### 3 Political Stability and Confidence in Economic Policy

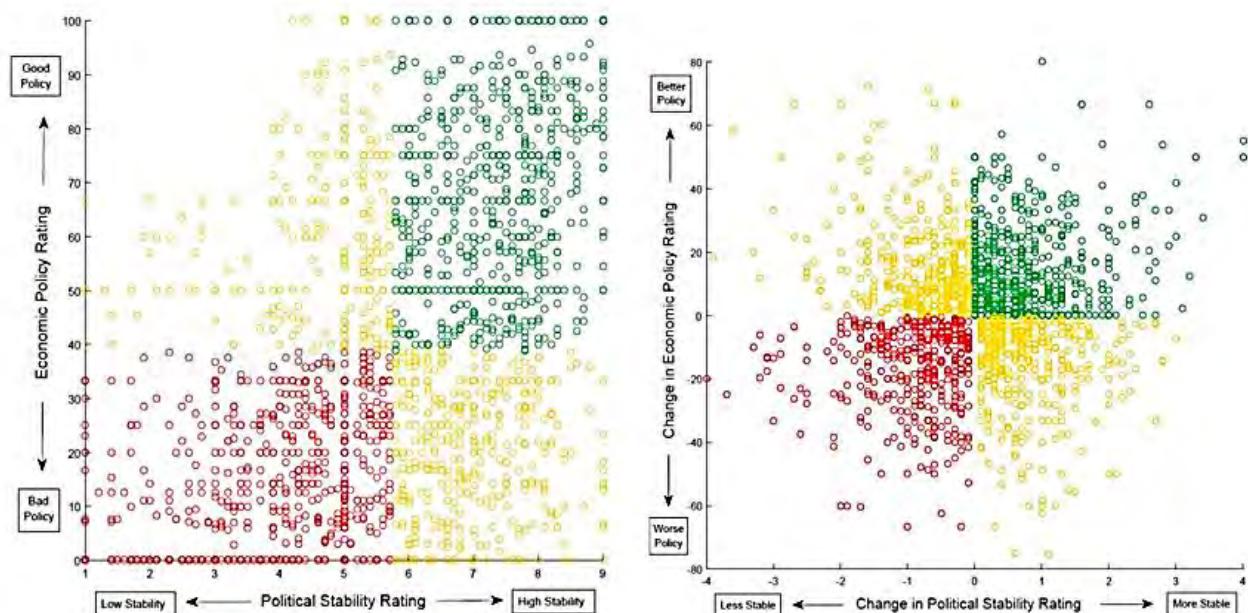
The work of Douglass C. North (1991) on institutions as “the rules of the game in a society or, more formally, the humanly devised constraints that shape human interaction” classifies rules into “political rules, economic rules, and contracts”.<sup>7</sup> Contracts are firm specific, potentially giving rise to idiosyncratic risk, but political and economic policy rules create systematic risks that should be priced in financial markets. North argues for separating the analysis of rules from economic policy choices.

Recently, Gala, Pagliardi, and Zenios (2018) studied the differential effects of politics and policy variables, using two variables from the Ifo World Economic Survey:

- i. The importance of political stability on the climate for foreign investors
- ii. The importance of confidence in the government's economic policy for the economy

A plot of these data for a sample of 42 countries over the period Jan. 1992 to Dec. 2016, reveals that politics and policy are not in tandem, see Figure 1. Countries with highly rated policies and low rated politics, and vice versa, are not just a few isolated cases, but instead, we observe them often, and for good reasons.

For instance, a country may be on a specific policy path no matter which party wins the elections, as was the case of the Greek fiscal adjustment program, implemented under a liberal government, a coalition of liberals with the socialists, and the radical left. Alternatively, a stable political system may face uncertainty in the economic policy agenda of a coalition government, as was the case after the 2018 German elections.



(Source: Gala, Pagliardi, and Zenios (2018) using the WES country ratings. Political stability is scaled from 1 to 9, and economic policy from 0 to 100, with higher scores denoting more stability and higher confidence.)

*Figure 1. Ratings of political stability and economic policy for 42 economies during 1992-2016.*

Political stability and confidence in economic policy have material differential effects on the financial markets. Gala, Pagliardi, and Zenios find that politics and policy have economically large and statistically significant impacts on returns, with their risk premia being almost additive. They document that international equity investment strategies that exploit the politics-policy predictability generate abnormal returns up to 14% p.a. Significant political premia are also observed for bond yields, exchange rates, and CDS spreads.<sup>8</sup>

<sup>7</sup> North was a co-recipient of the 1993 Nobel Memorial Prize in Economic Sciences for “applying economic theory and quantitative methods in order to explain economic and institutional change”.

<sup>8</sup> Based on ongoing work of one of us with Giovanni Pagliardi (BI, the Norwegian Business School) from a draft paper on “Political risks everywhere”.

Countries in the top and bottom terciles of political stability and confidence in economic policy, differ by 2.84% in their annual sovereign bond yields, 205bp in CDS spreads, and 13% in annual currency depreciations. Similarly, a country that improves its political stability ranking by 1 unit sees a reduction of its refinancing rates by 0.7%, with a reduction of 0.2% after a comparable improvement in economic policy confidence.<sup>9</sup> If Greece, for instance, would see its political stability downrated by 4 units on the WES scale, with its economic policy confidence remaining at zero (as it happened after the second bailout in 2012), its borrowing costs would deteriorate by 2.8%. *These changes in refinancing costs are attributed to the political risk factors, since the estimation regressions control for growth, unemployment, and debt. This allows us to incorporate political effects in a DSA model.*

We will use the model of Zenios et al. (2019) to optimize the sovereign debt issuance and ensure non-increasing debt dynamics (or reasonably fast-decreasing dynamics, for high debt countries). The model incorporates the critical dynamics of debt stock and debt flow (i.e., gross financing needs) into an optimizing problem, to minimize tail risk under macroeconomic, fiscal, and financial uncertainty, accounting for the feedback between debt stock and refinancing rates. Tails are the extreme quantiles (typically at the 0.95 level) of the stock and flow dynamics, and for sustainability they must remain within some thresholds. Uncertainty is represented using *scenario trees*, building on a long tradition of multi-period stochastic models that find numerous applications in the risk management of financial institutions.<sup>10</sup>

#### **4 Short-term political ambiguity and long-term uncertainty**

To incorporate political risks in DSA we are faced with the problem of uncertainty specification, which has been daunting economists for a very long time. Arrow (1951) distinguished *risk within a model*, where uncertainty is about outcomes that can be fully specified by a probabilistic model, and *ambiguity among models*, where we are uncertain about which model to use to assign probabilities. Knight (1921) had studies ambiguity three decades earlier, calling “risk” those situations where uncertainty could be described with a probabilistic model, and “true uncertainty” situations with partial ignorance about potential outcomes and their likelihood. To account for political risks in DSA we need to account for short-term ambiguity (i.e., which government wins the election, what policies will they institute, will a country be able to follow an adjustment program?), and the long-term volatility towards a well estimated expected future state, if we think that such an equilibrium state exists. In general, it is not possible to estimate reliably an election outcome or what policy a new government will follow. And even if this were the case, we are not faced with a repeated game, to work with expectations (or quantiles) over many repetitions.

How to incorporate both ambiguity about events in the short run and uncertainty about long-term trends? We adopt narrative scenarios for variables with ambiguous immediate outcomes, to see what the bad outcomes may be, and *calibrate probabilistic scenarios for long-run uncertainty to estimate appropriate risk metrics. Under ambiguity we run the DSA model for a range of plausible values for the critical variables that are affected by political event. For the long-run risks we calibrate scenarios of economic, fiscal, and financial variables, accounting for political effects. With this approach we identify values of ambiguous variables with high probability of bad outcomes, so that they can be avoided. The result is a comprehensive heatmap of high-risk debt dynamics, with quantile optimization for those aspects of the problem that are amenable to scenario calibration, and identification of narrative scenarios with bad outcomes that must be avoided, for the ambiguous aspects.*

#### **5 Political risks and the “red shift” in debt sustainability analysis**

Italy provides an interesting case study for a proof-of-concept. We extend our DSA model to evaluate the 2019 budget agreement between the Italian government and the EC. We assess if Italy can stay on a non-increasing debt path with gross financing needs below an IMF-specified threshold of 20% of GDP, and demonstrate the material effects of political risks. (Our assessment criteria are less stringent than those of official DSA.)

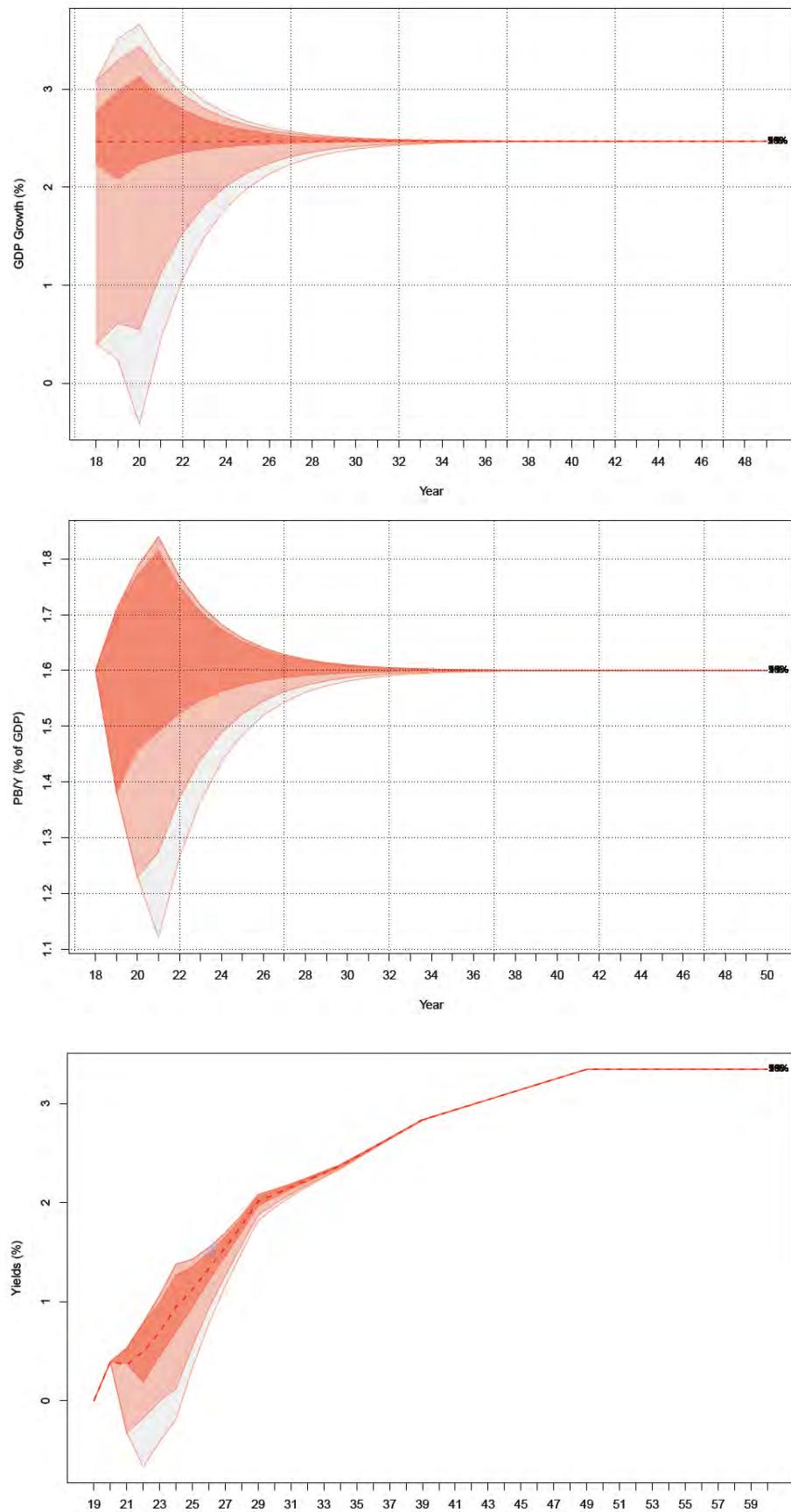
We start with a scenario tree of GDP growth, primary balance, and the risk-free rate of eurozone 5-year AAA rated sovereigns, but without political variables. The tree is calibrated to Italy’s conditions and observed market data, using historical volatilities and correlations. Fan chart samples from the tree are illustrated in Figure . To the scenarios of risk-free rates the model adds premia capturing the response of borrowing rates to debt levels.<sup>11</sup>

---

<sup>9</sup> These estimates were obtained running Fama-MacBeth cross-section regressions on the panel data in our sample of the yields of 10-year government bonds on the ratings of political stability and economic policy confidence.

<sup>10</sup> See, for instance, Mulvey, J. and W.T. Ziemba (eds.), Worldwide Asset and Liability Modelling, Cambridge, UK: Cambridge University Press, 1991, and S.A. Zenios. Practical Financial Optimization. Decision making for financial engineers. Malden, MA: Blackwell-Wiley Finance, 2007.

<sup>11</sup> The endogenous response of interest rates to debt levels and calibration to eurozone data, are in section 4.2 of Zenios et al. (2019).

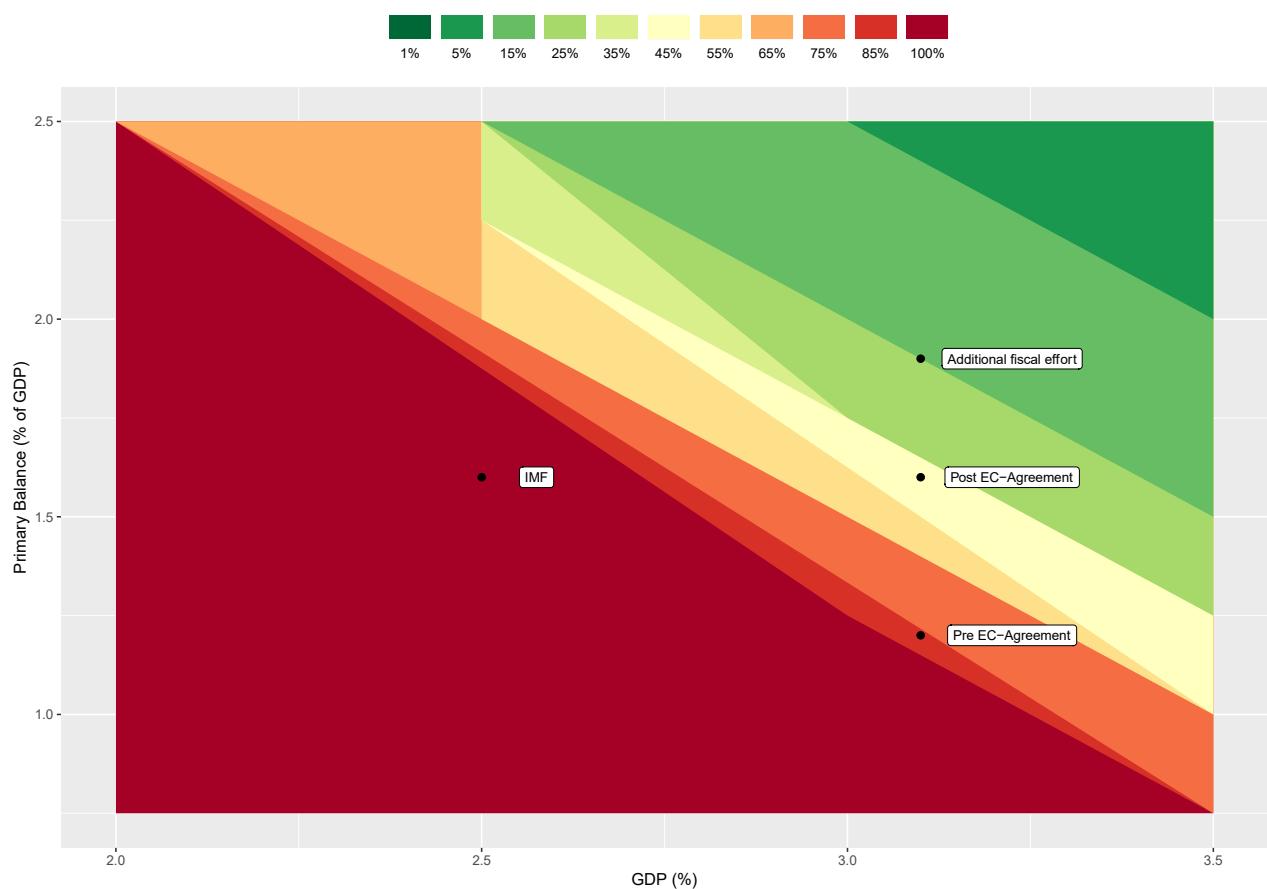


(Source: Calibration from Zenios et al. 2019, Figure C.1.)

Figure 2 Scenarios for (a) macroeconomic, (b) fiscal, (c) financial variables for Italy.

A significant short term political risk for DSA is the fiscal stance of the new government following the 2018 elections. We parametrically change growth and primary balance projections to cover plausible outcomes, and evaluate, using the calibrated long-term trees, the probability that debt stock and gross financing needs will stay within the thresholds. The result is a heatmap that shows the likelihood of debt dynamics remaining within the thresholds for a wide range of the ambiguous variables. We use the model to draw the heatmap and assess the Italian debt dynamics under three narrative scenarios: (i) no policy change, (ii) the new Italian government achieves its growth and surplus projections, and (iii) Italy reaches the targets of the negotiated EC agreement. For each narrative scenario we will assess if its outcomes violate the thresholds, and, therefore, must be avoided.

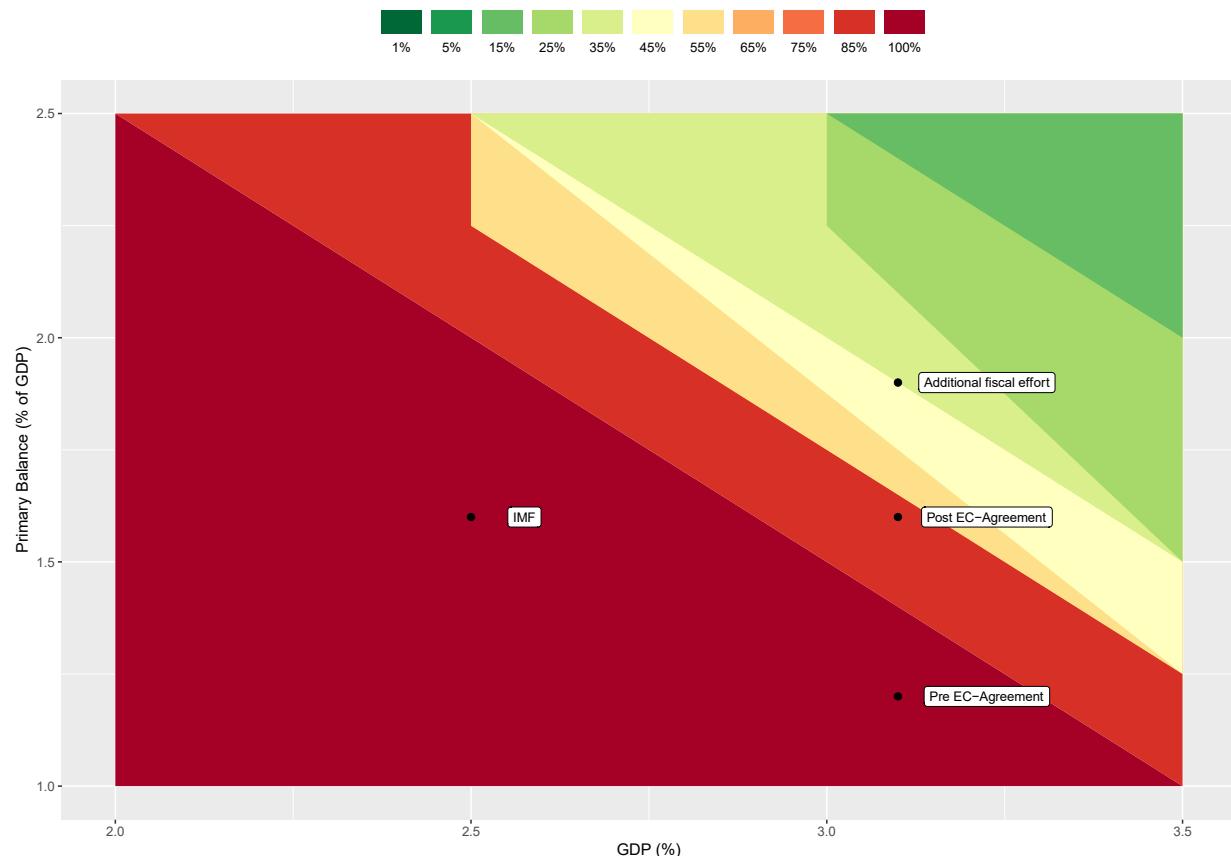
Figure 3 shows the heat map, with dark GREEN denoting extremely low probability (0.01) of unsustainable dynamics, and RED denoting very high probability (0.85). Note that for a wide range of combinations of GDP growth and primary balance the dynamics are unsustainable with very high probability. The country is clearly vulnerable, a point which was also emphasized in Sapir (2018) and for which there exists currently a broad consensus. On this map we place our narrative scenarios: "IMF" denotes projections from the IMF World Economic Outlook report for 2018, and under our model calibration and without any change in policy the debt dynamics are unsustainable with very high probability. "Pre EC-agreement" corresponds to the Italian government targets, and improves upon the current policy but it is still in the RED zone. "Post EC-Agreement" presents further improvements, shifting the country to the YELLOW zone, with probability 0.55 for sustainable dynamics. The sagacity of a policy with 0.55 chance of achieving its objectives is questionable, and additional fiscal effort is needed to increase the probability of remaining within the thresholds to 0.85 (GREEN). Using the model we estimate that a total fiscal effort of 3.5% of GDP over twelve years, capped at 0.3% p.a., Italy can reach this target. This finding is in agreement with Sapir (2018) that Italy should have been running consistently higher primary surplus to avoid finding itself in its current predicament, although our estimates for the extra effort are lower.



(Source: Authors calculations using the model of Zenios et al. (2019). Color-coded are the probabilities of increasing debt stock or gross financing needs violating the threshold over the risk horizon, for different combinations of primary balance and long-term growth. Also shown are results with the 2018 IMF projections for the Italian economy, the projections of the Italian government (Pre EC-agreement), and the projections agreed with the EC (Post EC-agreement). With additional fiscal effort of about 3.5% of GDP over the next twelve years the country enters the GREEN zone).

*Figure 3. Probability of the Italian debt violating thresholds on debt stock or gross financing needs.*

We now incorporate long-term political risks. We generate a new scenario tree with political stability and economic policy confidence state variables, calibrated to the country's volatile political variables around estimated long-term trends. To calibrate the political state variables we assume that they converge long-term to their historical averages of 4.5 for stability and 15.5 for policy. We also estimate volatilities from the historical ratings for Italy, namely a standard deviation of 1 for stability and 11 for economic policy confidence. The political variables are correlated with growth, primary balance, and interest rates, with historical correlations from -0.44 to 0.75, respectively. We now have a calibrated tree that accounts for political risk factors, with fan charts of the political variables qualitatively similar to those in *Figure 2*. Regression estimates of the bond yield sensitivities to these factors are then added to the scenarios of refinancing costs, adjusted with the endogenous debt risk premium. We re-run the model including political risk premia and redraw the heatmap in *Figure 4*.



(Source: Authors calculations as in the caption of Figure)

*Figure 4. Probability of the Italian debt violating thresholds on stock or gross financing needs under political risks.*

Comparing with *Figure 3* we note a marked “red shift” of the area with high probability of unsustainable dynamics. More combinations of growth and primary surplus have high probability of violating the thresholds. Under our model, the agreement with the EC, which was estimated to have a slightly better than 0.50 chance of success, fails with probability 0.85 when accounting for political risks. The additional fiscal effort that restores sustainability with probability 0.85, is now borderline light GREEN with 0.65 to 0.55 chance of success. Clearly, ignoring the political risks can lead to excessive optimism and wrong decisions.

## 6 Conclusions

Political risks can have material effects on sovereign borrowing costs. However, current debt sustainability analyses by the major international institutions ---IMF, ECB, and ESM--- treat non-quantitatively this significant risk factor. We argue that more rigor is both possible and essential, to incorporate political risks in DSA. We have also argued that DSA should go beyond testing unsustainable dynamics for distressed countries, to Debt Vulnerabilities Analysis for any country.

We show how a model can be extended to account for political risks. The short-term political risks cannot be treated probabilistically and we resort to narrative scenarios. Long-term political risks can be calibrated probabilistically, and we use recent advances in estimating political stability and economic policy confidence premia for the calibration. We combine the

narrative scenarios with the calibrated probabilistic scenarios to generate a comprehensive heatmap, and demonstrate using Italy as a case study. We assess the 2019 budget agreement between the Italian government and the EC, and document a significant “red shift” in the vulnerabilities of Italy when accounting for political risks.

Andrea Consiglio and Stavros Zenios

## References

- Arrow, K. J. Alternative approaches to the theory of choice in risk-taking situations. *Econometrica*, 19:404–437, 1951.
- Ben-Haim , Y. and M. Demertzis , Decision Making in Times of Knightian Uncertainty: An Info-Gap Perspective, *Economics*, 10:2016-2023, 2016.
- Bittlingmayer, G. Output, stock volatility, and political uncertainty in a natural experiment: Germany, 1880-1940, *The Journal of Finance*, 53:2243-2257, 1998.
- Bouabdalah, O., Checherita-Westphal, C., Warmedinger, T., de Stefani, R., Drudi, F., Setzer, R., and Westphal, A. Debt sustainability analysis for euro area sovereigns: a methodological framework, European Central Bank, Occasional Paper Series No. 185, 2017.
- Gala, V., G. Pagliardi, and S.A. Zenios, International politics and policy risk factors, Working Paper, 2018. ([https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3242300](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3242300))
- IMF, Staff Guidance Note for Public Debt Sustainability Analysis in Market-Access Countries, International Monetary Fund, Washington, D.C., 2013.
- Knight, F. H., Risk, Uncertainty, and Profit. Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Company, 1921.
- North, D. C., Institutions, Institutional Change, and Economic Performance, Cambridge, UK: Cambridge University Press, 1998.
- Sapir, A. High public debt in euro-area countries: comparing Belgium and Italy. Policy Contribution No. 15, September 2018.
- Schattschneider, E. Politics, Pressures and the Tariff, New York, NY: Prentice-Hall, 1935.
- Zenios, S.A., A. Consiglio, M. Athanasopoulou, E. Moshammer, A. Gavilan, and A. Erce. Risk management for sovereign financing with sustainability conditions. Globalization Institute Working Paper 367, Federal Reserve Bank of Dallas, 2019. (<https://www.dallasfed.org/~media/documents/institute/wpapers/2019/0367.pdf>)

# **External fraud detection through big data: towards a pro-active operational risk management**

di Giacomo Petrini

## **Abstract**

In anni recenti, l'utilizzo di reti neurali artificiali ha condotto alla formulazione di modelli econometrici euristici, la forma strutturale dei quali evolve in tempo reale, man mano che i dati disponibili crescono nel corso del tempo. Tali approcci hanno beneficiato anche della crescita esponenziale delle capacità tecnologiche di calcolo, nonché della crescente disponibilità di dati pubblicamente disponibili.

In particolare, il recente sviluppo di reti neurali artificiali auto-adattive consente di disporre di modelli econometrici, la specificazione dei quali evolve in tempo reale, che possono essere tra l'altro utilizzati per identificare in tempo continuo eventi che si realizzano raramente e che mutano nel corso del tempo; per tali ragioni, le reti neurali auto-adattive possono essere impiegate dalle istituzioni finanziarie anche per identificare gli stream comportamentali associati ad eventi esterni di frode, offrendo il vantaggio di consentire un rapido riconoscimento di nuove fattispecie non appena queste si manifestano, senza tuttavia tralasciare l'identificazione di eventi pregressi.

Le istituzioni finanziarie sono infatti particolarmente esposte ad eventi esterni di frode, che di norma presentano la peculiarità di accadere in tempi ristretti e che sempre più, negli ultimi anni, si relazionano con la crescente evoluzione tecnologica, ad esempio per quanto concerne il sistema dei pagamenti.

Pertanto, un efficace ed efficiente sistema di identificazione e gestione di tali eventi richiede sia velocità nell'esecuzione dei controlli, sia l'esistenza di processi e procedure efficaci per l'adozione di idonee azioni di mitigazione.

Le funzioni di controllo interno, specificamente Risk Management e Internal Audit, possono pertanto rappresentare gli attori primari nel guidare le attività di disegno, sviluppo e gestione di moderne soluzioni di identificazione delle frodi esterne, che consentano di identificare velocemente nuove tipologie di frode al loro primo manifestarsi e di definire altrettanto rapidamente azioni di mitigazione efficaci, in modo da ridurne gli impatti economici attuali e prospettici, che potrebbero rappresentare rischi per la profitabilità e la reputazione di una istituzione finanziaria.

## **1 An introduction to artificial neural networks**

In recent years, the exponential growth of calculation capacity has enabled the use both of longer time series and of more and more variables; the use of artificial neural networks has given rise to econometric models, evolving on a real-time basis as available data grow over time.

These artificial neural networks are particularly useful to make forecasts respecting a data universe that changes and/or grows quickly over time, particularly where the foreseen events are rare ones.

Generally speaking, an artificial neural network (ANN) is “*a massively parallel combination of [connected] simple processing unit which can acquire knowledge from environment through a learning process and store the knowledge in its connections*”<sup>1</sup>; these connections sometimes are called “synapses”, and the learning process is usually defined as the modification of synaptic weights to capture information. Indeed, an ANN requires three more basic elements to work: (i) a summing function, to combine input with respect to corresponding synaptic weights, (ii) an activation function, to produce a computed output to be compared with environmental evidences, and (iii) a feedback function, to modify backward the synaptic weights by reason of the difference between the computed output and the environmental evidences.

Thus, an ANN is basically an iterative computational approach that employs a dataset of explanatory variables to fit given environmental evidences in a multidimensional, unconstrained data universe: “*ANNs are distributed, adaptive, generally nonlinear learning machines built from many different processing elements (PEs). Each PE receives connections from other PEs and/or itself. The interconnectivity defines the topology. The signals flowing on the connections are scaled by adjustable parameters called weights. The PEs sum all these contributions and produce an output that is a nonlinear (static) function of the sum. The PEs' outputs become either system outputs or are sent to the same or other PEs.*”<sup>2</sup>

Therefore, the learning process of an ANN is a closed, dynamic computation. It has a beginning, once the four above mentioned basic elements have been created and once a dataset of input and output data has been prepared (so-called training dataset); it has a dynamic learning phase, during which the synaptic weights are iteratively modified by reason of the difference between the computed output and the environmental evidences; and it has an end, once the feedback function does not produce any further modification of the synaptic weights.

Thus, the ANN learning process outputs a closed (training-data driven) specification, i.e. a stable relation between input and output variables. This closed specification may then be applied to data, other than the training dataset, to obtain a model that explains environmental evidences.

<sup>1</sup> See Haykin S.; “Neural Networks: A Comprehensive Foundation”, Prentice Hall, New Jersey, 1999.

<sup>2</sup> See Rojas R.; “Neural Networks: A Systematic Introduction”, Springer-Verlag Berlin Heidelberg, 1996.

A new learning process is necessary whenever the computed output no longer represent an adequate fitting of the environmental evidences. For example, this happens when the analysis of the confusion matrix shows a declining capacity of correct classification over time.

Due to the exponential growth of calculation capacity, most recent developments allow to obtain “adaptive ANNs”, as the so-called Self-Organizing Maps (SOMs), or Kohonen Maps: SOMs can modify their specification on an ongoing basis, while the training dataset grows over time with the inclusion of the most recent data.

Thus, they can be applied to a dataset that grows in size over time, as they produce a dynamic specification that evolves over time. At every moment, if the dataset does not change, a SOM can calculate the best-fitting specification in relation to the given dataset: this specification is usually called “a map”.

Therefore, a SOM basically is “*a class of neural-network algorithms in the unsupervised-learning category*”<sup>3</sup> that is trained through unsupervised learning to produce a low-dimensional, discretized representation of the input space; SOMs operate in two ways: training and mapping: “training” builds the map using input examples, while “mapping” automatically classifies a new input vector.

In the SOM, the goal of learning is to cause different parts of the network responding similarly to certain input data. The synaptic weights are initialized either to small random values or sampled evenly from the subspace spanned by the two largest principal component eigenvectors.

The training uses competitive learning: when a training example is fed to the network, its Euclidean distance to all weight vectors is computed, and the PE whose weight vector is most similar to the input is called the “Best Matching Unit”, or BMU. The weights of the BMUs and PEs, in the SOM grid, are adjusted towards the input vector. The magnitude of the change (or learning rate) decreases with time and with the grid-distance from the BMU.

Thus, SOMs differ from other ANNs because they apply competitive learning opposed to error-correction learning (such as back-propagation with gradient descent). They use a neighborhood function to preserve the topological properties of the input space, even if a SOM (unlike an ANN) can evolve its own topology over time.

To a greater level of detail, in the SOM algorithm, a Hebb-like learning rule is usually used, with time decreasing learning parameters. For all the PEs, the learning rate and the magnitude of its change decrease with time during the training phase, as the map stabilizes and learns the topographic map of the training dataset. At the final step, the learning-rate parameter usually has a very small value and so does the neighborhood function.

Therefore, the SOM algorithm cannot learn with adequate speed the new environmental evidences that may be different in statistical characteristics from the previously learnt training data: the time-decreasing learning rate and neighborhood function of the basic SOM algorithm reduce its capability to adapt weights for a varied environment.

A specific extension of the basic SOM is represented by the Time-Adaptive Self-Organizing Map (TASOM): it employs adaptive learning rates and neighborhood functions as its learning parameters, and every PE has its own learning rate and neighborhood size. For each new input vector, the neighborhood size and learning rate of the winning PE and the learning rates of its neighboring PEs are updated.<sup>4</sup>

Therefore, the learning parameters of every PE are time-independent and they change their values with the conditions of the incoming data, not with the elapse of time. Moreover, the learning parameters of each PE are adjusted independently for each PE, as the learning rate of each PE follows the values of a function of distance between the input PE and its synaptic weight. A similar updating rule is used to automatically adjust the width of the neighborhood function of each PE. The width of each output PE follows the distance between the PE synaptic weight and the weight of its neighboring PEs.

The operation of the TASOM algorithm may be summarized in eight steps:

1. Initialization: as in the case of a SOM, the synaptic weights are initialized either to small random values or sampled evenly from the subspace spanned by the two largest principal component eigenvectors;
2. Sampling: as in the case of a SOM, the training dataset is fed to the TASOM;
3. Similarity matching: the best-matching (or winning) PE is found, using the minimum-distance Euclidean norm as the matching measure;
4. Updating the neighborhood width: the neighborhood width of the winning PE is adjusted, while the neighborhood widths of the other PEs do not change;
5. Updating the learning-rate parameters: the learning rate parameters of all PEs in the network are adjusted independently for each PE;
6. Updating the synaptic weights: the synaptic weights of all PEs in the network are updated;
7. Updating the scaling values: the scaling values of all output PEs in the network are updated. A TASOM normalizes all distance calculations so as each distance calculation in the network algorithm is normalized by a scaling vector, composed of standard deviations of input vectors’ components;
8. Continuation: the procedure restarts from step 2.

<sup>3</sup> See Kohonen, Kaski, Somervuo, Lagus, Oja, Paatero, 2004.

<sup>4</sup> See Shah-Hosseini, Safabakhsh, 2002, and Shah-Hosseini, Safabakhsh, 2003.

SOMs and TASOMs are particularly useful to make forecasts with respect to a data universe that changes and/or grows quickly over time, particularly where the events to be foreseen are rare ones.

## 2 Detecting external frauds in the financial sector with Data Analytics

A recent line of research deals with the application of SOMs and Data Analytics to the detection of external frauds.<sup>5</sup> External fraud events in the financial sector are usually not so frequent events, even if they can imply significant actual and perspective economic impacts, as they involve significant financial risks which may threaten profitability, and the image of a company.

Moreover, external frauds may assume several specific forms (e.g. credit cards forgery, identity theft, phishing, skimming, mobile SIM swapping, account hijacking, check-kiting schemes), but events usually happen in a short period of time and the technology side is more and more involved as payment systems evolve towards new technologies.

Thus, external fraud detection has to be fast to be effective, and usually the detection process requires the acquisition of several real-time raw data<sup>6</sup> to be managed under time pressure.

In these circumstances, the development of appropriate IT applications plays a central role in the creation of an effective detection process, as this issue can no longer be manually performed because it requires the use of data analysis tools and programs to be applied to large volumes of data.

Conventional fraud detection is sometimes ineffective, for example in online fraud, because online fraud is very dynamic and also technology development for conducting fraud is very dynamic and constantly changing.

Therefore, most recent evolutions are turning towards processes of continuous monitoring to identify potentially fraudulent anomalies in the data stream or in behavioral patterns, and to apply the result to prevent external frauds:<sup>7</sup> anomalies are identified starting from the activities or behavior typically observed in a single user (e.g. geospatial localization of the login site; identification of the computer model used to login; identification of the internet service provider used to login; activities performed after login; etc.), as this behavioral pattern can be used to establish an online behavior model of the user which is very specific and unique to him.<sup>8</sup>

This makes fraud easier to detect, because typically the fraudster does not know how the user behaves online, so it may be very difficult for the fraudster to appear like the account owner.

Such a behavioural analysis can also be integrated using other observable parameters of the session (e.g., IP address, HTTP header information, page views, etc.):

- the IP address provides an estimate of geospatial location, and information like country, city, network block, internet service provider;
- the HTTP header provides information of the operating system, user agent string, referrer string and browser type of the computer used to login;
- session time, session timeslot and session duration;
- amount, counterparty and details of the transaction, if any;
- etc.

On the analytical market, there is a wide spectrum of specialized tools that can support and enhance the antifraud activity. Anyway, their use is still limited, even if a proactive data monitoring/analysis remains one of the most effective tool for antifraud control, to reduce fraud losses and fraud scheme duration.<sup>9</sup>

Most recent continuous monitoring processes use SOMs and Data Analytics, combining specific methods and techniques from computer science, mathematical sciences, statistical, economic, psychology, law and other cognitive sciences. Usually, these processes apply operational analysis on short term, exploiting data and current information to detect fraud events with maximum efficiency.

Data analysis, in operational form, becomes also a tool for improving workplace conditions: most manual activities are avoided, especially for operations that require to process a significant amount of data.

<sup>5</sup> For a decade review from 2004 to 2015, see Albashrawi, Mousa, "Detecting Financial Fraud Using Data Mining Techniques: A Decade Review from 2004 to 2015", Journal of Data Science, Vol. 14, 2016.

<sup>6</sup> Nowadays text data, pictures, audio, video etc. comes from everywhere, from many different sources like contracts, customer interactions, call centers, social media, phones, emails, faxes, and others.

<sup>7</sup> A significant part of the involved data may also have to be stored, should it be later used in direct investigations. The Internal Audit should also access these data to make recommendations to improve the control activities.

<sup>8</sup> For present purposes, a "behavioral pattern" may be defined as a time sequence of actions, each of which produces measurable effects; the sequence always happens in the same chronological order, so that the chronological sequence of actions is relevant.

<sup>9</sup> For a review of the application of ANNs and SOMs to credit card fraud detection, see: Seja, K. R.; Zareapoor, Masoumeh; "FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining", Hindawi Publishing Corporation, The Scientific World Journal, Vol. 2014; Akhilomen, John, "Data Mining Application for Cyber Credit-card Fraud Detection System", Proceedings of the World Congress on Engineering, Vol. 3, 2013.

In daily activities, the main role of operational analysis is to help detecting and fighting illegal activities, examining: (i) the links between suspects, (ii) their characteristics (director/and indirect subordination relations, positions in the hierarchy of the group, key positions that impact decision-making etc.), (iii) the movement of goods, money or other valuables, (iv) the way of communication (email, social networking), (v) the sequence of certain events, (vi) the modus operandi etc.

For example:

- sub (i) and (ii), the examination of the links between suspects and their characteristics may be relevant to identify “payment transactions made as a result of the payer being manipulated by the fraudster to issue a payment order, or to give the instruction to do so to the payment service provider, in good-faith, to a payment account it believes belongs to a legitimate payee (s.c. ‘manipulation of the payer’)<sup>10</sup>”;
- sub (iii), the examination of the movement of goods, money or other valuables may be relevant to identify trade-based money laundering schemes, i.e. the process of disguising the proceeds of crime and moving value through the use of trade transactions in an attempt to legitimize their illicit origins; in practice, this can be achieved through the misrepresentation of the price, quantity or quality of imports or exports;
- sub (iv), the examination of the way of communication may help to discover potential illegal actor networks cases, where fraudsters pretend to be part of, or authorised by, a company and in doing so attempt to carry out ‘scams’ or frauds; these may involve offering a large sum of money to the recipient such as an inheritance or unclaimed bank funds, either in return for paying a modest sum (often referred to as transaction fees or brokering costs) or requiring the recipient’s identity details;
- sub (v), the examination of the sequence of certain events may be relevant to identify the anomalies of the current spending pattern of a customer, with respect to his or her past spending behavior, so as to properly intercept new transactions that deviate from the learned profiles;
- sub (vi), the examination of the modus operandi may help to identify payment frauds through unauthorized access to mobile phone and/or online transactions, and to identify the precautions that need to be taken.

The success of this approach typically lies in the quality and variety of data sources.

Data mining, as an analytic process, is designed to explore data and to extract information from data sets, in order to discover patterns and relations.

It can be defined as “*the nontrivial extraction of implicit, previously unknown, and potentially useful information from data*”<sup>11</sup>, or “*the science of extracting useful information from large data sets or databases*”<sup>12</sup>.

Data mining shall include (but not be limited to):

- a. analysis of the data-text, known as exploitation of data such as text, or "text mining"; it refers to the process of extraction of knowledge from documents, because information can be mostly found in text format, and it is considered one of the most important area in the database system;
- b. geospatial analysis and visual analysis; they are important to understand the relevance of the location where events happened, to determine and discover patterns in fraud behavior.

In fact, successful implementation of an antifraud analytical system highly depends on the manner of retrieving data from a variety of sources, considering that most of them have different formats.

The data collected should be interpreted in the same way, using the same techniques and the same methodology, so that creation of data bases to be homogeneous.

### **3 A Self-Organizing Map for external fraud detection**

We assume that nothing is known about external frauds, i.e. we don't have any aprioristic theory explaining the phenomenon. We also assume that we have a (scarce) history regarding external fraud events that hit our customers in the past (m events), and a (wider) history regarding legal transactions performed by the same customers (n events, n>m). Each transaction (legal or not) is accompanied by the time of its occurrence. Moreover, before each transaction occurrence, we can retrieve:

- a set  $\Gamma$  of unstructured private data regarding the behavior of each customer of the bank;

<sup>10</sup> See European Banking Authority, “EBA/GL/2020/01 - Guidelines amending Guidelines EBA/GL/2018/05 on fraud reporting under the Payment Services Directive (PSD2)”, January 22, 2020.

<sup>11</sup> See Frawley, William J.; Piatetsky-Shapiro, Gregory; Matheus, Christopher J.; “Knowledge Discovery in Databases: An Overview”.

<sup>12</sup> See Hand David J.; Mannila, Heikki; Smyth, Padhraic; “Principles of Data Mining”.

- (through the access to customers' publicly available data, e.g. public profiles of their social media) a set  $\Lambda$  of unstructured public data.

The only variable that can let us organize  $\Gamma$  and  $\Lambda$  is the time: every piece of data embed the time of its occurrence, so that  $\Gamma$  and  $\Lambda$  can be ordered by the time of occurrence.

Thus, the dependent variable is represented by a vector  $\Psi$  composed of  $(n+m)$  rows, where every row assume the value 1 in the case of a fraud event, and 0 in the case of a legal transaction. It is a question of finding those relations between (current and lagged) data included in  $\Gamma$  and  $\Lambda$  that define the multidimensional hyper-surface (not necessary a hyperplane) that divides the estimation space between fraud events and legal transactions, as every transaction can be related to the actual observable parameters (including time, IP address,<sup>13</sup> browser, operating system, HTTP header,<sup>14</sup> etc.) corresponding to an event.

Although this approach sounds similar to the one applied to estimate models for credit risk parameters, the main difference is that the scope is not to estimate a model for forecasting fraud events.

Rather, it is to discover those multidimensional behavioral patterns, across time, that can be associated unambiguously to fraud events. Sometimes, the solution may be quite easy, considering any particular activity carried out in the account while logged into the account, as for example:

- an ATM withdrawal recorded in a certain country, while geo-localization data embedded in a public post on a social media show that the customer was in another country at the same time of the withdrawal;
- a credit card payment made from a masked IP address, while behavioral data of our customer show that he never used IP-masking in the past;
- an online banking transaction made from an IP address of a certain country, while geo-localization data, embedded in a public post on a social media, show that the customer was in a faraway place at the same time of the transaction;
- an online banking transaction made from a different computer ID, and/or with a different HTTP header information.

Indeed, detection can prove a more difficult task for fraud events different from those related to the payment system. All of these concern more structured behavioral patterns, sometimes involving groups of connected customers.

Anyway, the exact time sequence of the fraudster actions becomes important using this approach. In fact, the analysis focuses on the kinds of things users generally do before a fraud event, thus detecting specific known fraud patterns.

For these reasons, a heuristic neural network approach seems preferable, as it does not necessarily require linear relations between dependent ( $\Psi$ ) and independent ( $\Gamma$  and  $\Lambda$ ) variables, neither it requires an aprioristic model identification<sup>15</sup>.

Instead, this approach lets the data identify those structural relations that can offer the better explanation of fraud events.

Furthermore, fraud events do not represent a close universe over time: both technology evolution and the adoption of effective mitigation techniques can modify the set of events to be detected, or because new types of external fraud arise, either because some kind of them are effectively mitigated<sup>16</sup>.

Thus, a TASOM can be applied for detection purposes of external fraud events, as it upgrades on an ongoing basis, while the training dataset grows over time with the inclusion of the most recent data.

In fact, as the adaptive capabilities of a SOM or a TASOM modify when the magnitude of the changes decreases with time, these algorithms (when developed recursively over time) could be particularly useful to identify the behavioral patterns of external frauds, compared to most recent data, without neglecting the peculiarities of most ancient data.

Thus, this approach can detect new kinds of frauds, even though these new frauds have never been seen before, because it is based on the user's behavior, while old fraudulent activities and events are still taken into account, with their specific known rules, patterns, and/or indicators.

This approach can be used to generate alerts or warnings for real-time events where applied to real time transactions. Every event starts with someone, either a user or a fraudster, initiating an observed event that includes, for example, someone logging in to the user's account and/or any activity taken during an online session (e.g., checking account balance, transferring funds between accounts, viewing account information, etc.).

The observed event may or may not be an online event. Each event includes or corresponds to one or more event parameters, collected by the set of unstructured private data. Event parameters are directly observable parameters of an event, or raw data that can be measured and/or observed. Examples of event parameters incorporate, but are not limited to:

<sup>13</sup> Usually, the IP address provides an estimate of location information like country, state, city, network block, and internet service provider.

<sup>14</sup> Usually, the HTTP header provides information of the operating system, user agent string, referrer string, and browser type of a computer used for an event.

<sup>15</sup> In econometrics, the identification problem consists in identifying both the exogenous and endogenous variables and the relation between them.

<sup>16</sup> For example, at the beginning of the century, the frauds involving cloning ATM cards were effectively mitigated through the introduction of micro-chipped cards, so that related external frauds were considerably reduced.

- network information that include parameters of the network by which an online event is occurring (e.g., IP address, and country, state, city are derived parameters derived from network information),
- user agent parameters (operating system and browser of device or computer used for the event),
- event or session time,
- time sequence of deposits and withdrawals and/or of other incoming and outgoing bank records,
- time sequence of the banking transactions and operations conducted from, to or via one or more bank accounts or by specified persons during a specified period.

#### **4 Potential contribution to Operational Risk Management (ORM) and Audit functions**

According to their respective attributions, both the ORM and the Internal Audit functions are in some way involved in the detection and monitoring processes of external frauds. On one hand, the Operational Risk Management (ORM) is liable with regard to monitor and measure over time the loss events, including external frauds, and to coordinate the design of appropriate and effective mitigation actions. On the other hand, the Internal Audit function is liable with regard to the evaluation of the continuous effectiveness of the internal control system.

External frauds events usually happen in a short period of time and the technology side is more and more involved as payment systems develop new technologies. Therefore, external fraud detection has to be fast to be effective, and usually the detection process needs to acquire several real-time raw data to be managed under time pressure.

In this respect, TASOMs can be effectively applied for detection purposes, because they can apply operational analysis and exploit data and current information to detect fraud events with maximum efficiency.

Moreover, they can keep memory of previously identified fraudulent behavioral patterns, while upgrading on an ongoing basis, as the training dataset grows over time with the inclusion of the most recent data; they can therefore detect new kinds of frauds, even though these new frauds have never been seen before, while old fraudulent activities and events are still taken into account.

Thus, at the design stage, the ORM function may liaise with the IT function to draw up the application architectures, considering properly that the IT applications should be strictly interlinked with transactional data and databases, rather than with summary databases. Moreover, the same ORM function may also liaise with the Organization function to define the whole process of external frauds identification and management, which should ensure high effectiveness to be achieved in a limited period of time.

On the other hand, still at the design stage, the Internal Audit function may review the project activities, both to get comfortable with the IT and organizational developments and to ensure that all the relevant information is acquired and treated in a proper way.

During the operational phase, both the above mentioned functions can analyze the SOM results to decide which action is needed: the ORM function may focus on high frequency and/or high impact external fraud events (to coordinate the design of appropriate and effective mitigation actions), while the Internal Audit function may focus on emerging external fraud schemes, to evaluate the continuous effectiveness of the internal control system and, where appropriate, to address suitable evolutionary actions. A similar approach could be applied more generally by the Audit function to other areas of interest for its activities. Examples of potential application could be, but not limited to:

- the detection and the prevention of internal fraud events; TASOMs could be applied to identify the fraudulent behavioral patterns of company's employees, and the application could employ also the accounting information recorded through internal front-end procedures and the detailed data contained inside the so-called "giornale di fondo";
- the identification of the behavioral patterns related to money laundering phenomena; TASOMs may be applied to identify the customers' behavioral patterns that are related to the concealment of the origins of illegally obtained money, typically by means of transfers involving foreign banks or legitimate businesses;<sup>17</sup>
- the accounting audit activities, to analyze accounting data and identify unsustainable forecasting accounting variable time-series (such as cash flows or earnings) and/or anomalous accounting entries.<sup>18</sup>

In the above mentioned areas of interest, the application of TASOMs represent an effective analytical instrument to quickly identify behavioral sequences that are related to given output phenomena; these behavioral patterns, once identified through the training dataset, give shape to the learning parameters that, from then on, may increase or decrease for adaptation to a changing environment with non-stationary input distributions.

---

<sup>17</sup> See for example the series "Quaderni dell'antiriciclaggio" published by the "Unità di Informazione Finanziaria per l'Italia" (s.c. UIF), which includes several examples of misconduct aimed at "washing" illegally obtained money; these behavioral schemes can be easily translated in a chronological set of actions inside a TASOM.

<sup>18</sup> See Kokina J.; Davenport T. H.; "The Emergence of Artificial Intelligence: How Automation is Changing Auditing".

Moreover, in the same areas of interest, it is relevant to keep memory of previously identified behavioral patterns, while the detection framework upgrades on an ongoing basis, as most recent data are included in the training dataset; TASOMs can vouch for that, taking into account old behavioral patterns, and they can also detect and identify new kinds of events, even though they have never been seen before.

Anyway, to a more general level, TASOMs and other ANNs could be employed both by Risk Management and Audit functions in many other fields, for example (but not limited to):

- the risk assessment activities: through the use of data extracted from real cases, the Audit function could identify higher-risk processes and evaluate the effectiveness of the internal control system.<sup>19</sup> Similarly, the Risk Management could employ external data to identify higher-risk credit or financial sub-portfolios, and to define suitable mitigation actions;
- the scenario analysis: ANNs could be used both to generate a plethora of stochastic scenarios and to identify the transmission mechanism to final control variables;
- the project risk management: ANNs could be used in the scheduling of complex projects, to identify hidden relation between single tasks and to help the project management; unlike the use of PERT charts, the use of ANNs could consider iterations and feedbacks, where tasks may have to be reworked, so as to take note of task iterations and iterative development processes.

## 5 Concluding remarks

Recent developments in econometric research can take advantage of both the exponential growth of calculation capacity and the growing availability of public data. Recently, new procedures are quickly establishing: the use of artificial neural networks is giving rise to econometric models that evolve on a real-time basis, as available data grow over time. These may be used to identify rare events at any particular time, on the basis of the information available until that moment.

SOMs and TASOMs are more and more used to identify the behavioral patterns of external frauds, with respect to most recent data, without neglecting the peculiarities of most ancient data.

Thus, this approach can detect new kinds of frauds, even though these new frauds have never been seen before, because it is based on the user's behavior. Furthermore, old fraudulent activities and events are still taken into account, with their specific known rules, patterns, and/or indicators.

Financial institutions are particularly exposed to external frauds: external fraud events usually happen in a short period of time and the technology side is more and more involved as payment systems evolve towards new technologies.

Therefore, external fraud detection has to be fast to be effective, and usually the detection process requires to acquire several real-time raw data to be managed under time pressure.

The internal control functions, primarily ORM and the Internal Audit, can be involved in the design, development and operational phases of any modern solutions that quickly identify and manage emerging external fraud schemes, so as to swiftly define and apply effective mitigation actions, and to reduce actual and perspective economic impacts which may threaten profitability and the image of a financial institution.

Giacomo Petrini

---

<sup>19</sup> See Ramamoorti, S.; Bailey, A. D. Jr; Traver, R. O.; "Risk Assessment in Internal Auditing\_ A Neural Network Approach".

## Bibliography

- AIFIRM – Associazione Italiana Financial Industry Risk Managers, “Position Paper N° 14 - Intelligenza Artificiale: l'applicazione di Machine Learning e Predictive Analytics nel Risk Management”, February 2019
- Akhilomen, John, “Data Mining Application for Cyber Credit-card Fraud Detection System”, Proceedings of the World Congress on Engineering, Vol. 3, 2013
- Albashrawi, Mousa, “Detecting Financial Fraud Using Data Mining Techniques: A Decade Review from 2004 to 2015”, Journal of Data Science, Vol. 14, 2016
- Baldwin, Amelia A.; Brown, Carol E.; Trinkle, Brad S.; “Opportunities for Artificial Intelligence Development in the Accounting Domain: The Case for Auditing”, Intelligent Systems in Accounting, Finance and Management, Vol. 14, 2006
- Banarescu, Adrian, “Detecting and Preventing Fraud with Data Analytics”, Procedia Economics and Finance, Vol. 32, 2015
- European Banking Authority, “EBA/GL/2020/01 - Guidelines amending Guidelines EBA/GL/2018/05 on fraud reporting under the Payment Services Directive (PSD2)”, January 22, 2020.
- Fletcher, H. Glancy; Surya, B. Yadav; “A computational model for financial reporting fraud detection”, Decision Support Systems, Vol. 50 (3), February 2011
- Frawley, William J.; Piatetsky-Shapiro, Gregory; Matheus, Christopher J.; “Knowledge Discovery in Databases: An Overview”, AI Magazine, Vol. 13, Number 3, 1992
- Hand David J.; Mannila, Heikki; Smyth, Padhraic; “Principles of Data Mining”, MIT Press, 2001
- Haykin S.; “Neural Networks: A Comprehensive Foundation”, Prentice Hall, New Jersey, 1999
- Kohonen, Teuvo, “Exploration of very large databases by self-organizing maps”, Proceedings of International Conference on Neural Networks (ICNN'97), June 1997
- Kohonen, Teuvo; Kaski, Samuel; Somervuo, Panu; Lagus, Krista; Oja, Merja; Paatero, Vesa; “Self-organizing map”, CIS Biennial Report 2002-2003, Ch. 8, 2004
- Kokina, Julia; Davenport, Thomas H.; “The Emergence of Artificial Intelligence: How Automation is Changing Auditing”, Journal of Emerging Technologies in Accounting, Spring 2017, Vol. 14, No. 1, pp. 115-122
- Ramamoorti, Sridhar; Bailey, Andrew D. Jr; Traver, Richard O.; “Risk Assessment in Internal Auditing: A Neural Network Approach”, International Journal of Intelligent Systems in Accounting, Finance & Management, Vol. 8, 1999
- Rojas R.; “Neural Networks: A Systematic Introduction”, Springer-Verlag Berlin Heidelberg, 1996.
- Seeja, K. R.; Zarepoor, Masoumeh; “FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining”, Hindawi Publishing Corporation, The Scientific World Journal, Vol. 2014
- Shah-Hosseini, Hamed; Safabakhsh, Reza; "TASOM: A New Time Adaptive Self-Organizing Map", IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, Vol. 33 (2), April 2003
- Shah-Hosseini, Hamed; Safabakhsh, Reza; "The Time Adaptive Self-Organizing Map for Distribution Estimation", International Journal of Engineering, Vol. 15, No. 1, February 2002
- Zhang, Bin; Zhou, Yi; Faloutsos, Christos; “Toward a Comprehensive Model in Internet Auction Fraud Detection”, Proceedings of the 41st Hawaii International Conference on System Sciences, 2008

# Nuove policy nazionali ed internazionale: possibili implicazioni sulle Banche Italiane

di Camillo Giliberto (MPS)<sup>1</sup>

## Abstract

The article proposes some reflections on the funding and bank credit policies of italian banks in the light of the considering originated by some factors although different in times, in the modalities and certainty of manifestation are expected to generate negative effects on the economic-financial system.

In an economic context with growth estimates revised downwards for the countries of the euro area (and not only), the new regulations introduced by the regulators appear likely to have a profound impact not only on the structure of the banking sector, but also on the configuration of other sectors, primarily the real economy.

Bank lending to the private sector has slightly weakened: lending to the firms are decreased. The increase in the cost of the funding recorded last year has so far been transmitted to a limited extent to the interest rates charged on loans to customers, also thanks to the high liquidity and the improvement of banking balance; however, there are signs of tightening conditions for access to credit.

## Premessa

Nel 2018, gli utili degli istituti di credito sono tornati a valori positivi (le prime otto banche italiane hanno registrato un utile netto complessivo superiore ai 9 miliardi di euro) grazie al calo degli accantonamenti e al miglioramento dell'efficienza operativa ma **la bassa redditività rappresenta ancora il principale fattore di rischio**.

In presenza di tassi di interesse negativi, i margini unitari delle banche rimarranno compressi almeno per tutto il 2019.

In questo quadro e in presenza delle limitazioni poste dalle nuove regole europee, il sistema bancario europeo si trova a fronteggiare uno scenario sempre più complesso.

Da un lato, le nuove norme imposte dai regulators internazionali (nuove norme sui Crediti, introduzione del Requisito MREL) potrebbero determinare sia una maggiore rigidità nella concessione del credito che un aumento del costo della raccolta che potrebbe essere tanto più ampio quanto maggiore è il rischio sovrano.

Dall'altro, l'incremento del costo di provvista, metterebbe ulteriormente a rischio i profitti degli istituti di credito (in particolare per le banche che non emettono obbligazioni in modo continuativo – Fonte Banca D'Italia) già insufficienti per remunerare adeguatamente i costi di capitale.

Nel presente lavoro si propongono alcune riflessioni sulle politiche del funding e del credito delle banche italiane, alla luce dei condizionamenti originati da tre fattori:

- i) aumento del premio al rischio sovrano;
- ii) nuove norme sui crediti anomali;
- iii) nuove regole MREL.

Quest'ultimi fattori, benché diversi nei tempi e nelle modalità, sono attesi generare effetti negativi sulla redditività che, nel primo trimestre 2019 (fonte banca d'Italia), ha registrato una riduzione rispetto al corrispondente periodo nell'anno precedente (Fonte Banca d'Italia), risentendo del calo delle commissioni nette pur continuando a beneficiare del decremento dei costi operativi e delle rettifiche su crediti.

Nei paragrafi successivi, vengono analizzati in modo peculiare gli aspetti più generali.

## 1 Crescita economica e rischio sovrano

L'attività economica dell'Area dell'Euro nel 2019 resta debole e soggetta a rischi al ribasso. L'inflazione rimane su valori contenuti. Sulla base delle proiezioni elaborate nel settembre del 2019 dalle banche centrali dell'Eurosistema, il PIL aumenterebbe **dell'1,1 % nel 2019, dell'1,2% nel 2020 e del 1,4% nel 2021** (grafico 1).

Le prospettive per l'eurozona presentano diversi rischi al ribasso, legate prevalentemente all'incertezza sulla gestione della Brexit e alla vulnerabilità nei mercati emergenti. Il PIL dell'area euro nel primo trimestre 2019 si è attestato allo 0,4% (tab. 1) rispetto allo 0,2% registrato nell'ultimo trimestre 2018.

Fra i maggiori paesi, l'attività ha accelerato in **Spagna, Germania e, in misura minore, in Italia, mentre ha appena rallentato in Francia**.

All'interno dell'Area Euro, nel primo trimestre 2019 il PIL italiano ha registrato un rialzo congiunturale dello 0,1% (tab. 1), dopo due trimestri consecutivi in cui il sistema economico italiano si era contratto (grafico 2) entrando in recessione tecnica.

<sup>1</sup> Le opinioni espresse dall'autore sono personali e non impegnano l'istituto di appartenenza. L'autore, pur rimanendo completamente responsabile dei contenuti dell'articolo desidera ringraziare la Dott.ssa Ilaria Andreani per l'attività di revisione dell'elaborato.

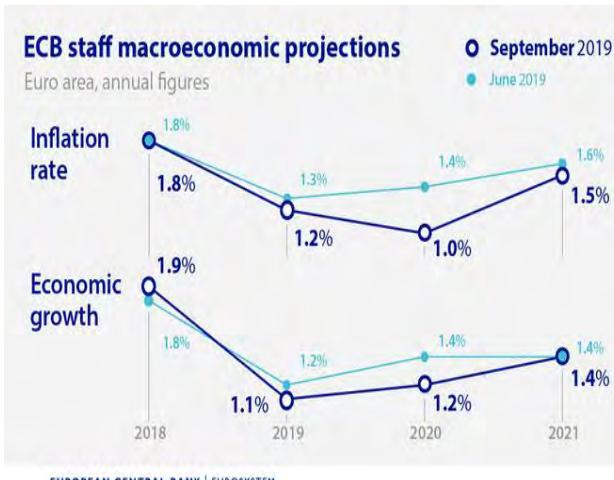
Nel secondo trimestre del 2019, secondo l'Istat, il Pil italiano sarebbe invece rimasto stazionario rispetto al trimestre precedente registrando una contrazione dello 0,1% rispetto allo stesso periodo del 2018.

Secondo proiezioni elaborate da Banca d'Italia, il PIL in Italia sarebbe pari allo 0,1 % nel 2019, allo 0,8% nel 2020 e all'1,0% nel 2021.

Il quadro è caratterizzato da un rallentamento degli investimenti e dalle esportazioni che risentirebbero della decelerazione del commercio mondiale.

A partire dalla seconda metà dell'anno in corso l'attività recupererebbe gradualmente, soprattutto grazie alla spesa delle famiglie e alle esportazioni.

Grafico 1- Crescita economica nell'area euro (Fonte BCE)



Tab. 1 – PIL e Inflazione Area Euro (Fonte Istat e Eurostat)

### Crescita del PIL e inflazione nell'area dell'euro (variazioni percentuali)

PAESI	Crescita del PIL		Inflazione	
	2018	2018 4° trim. (1)	2019 1° trim. (1)	2019 giugno (2)
Francia	1,7	0,4	0,3	1,4
Germania	1,4	0,0	0,4	1,3
Italia	0,9	-0,1	0,1	0,8
Spagna	2,6	0,6	0,7	0,6
Area dell'euro	1,9	0,2	0,4	1,2

Fonte: elaborazioni su statistiche nazionali e su dati Eurostat.

## 1.1 Rischio sovrano ed effetti sul sistema bancario

Nel corso del 2019, il profilo di funding e di liquidità delle banche italiane è tornato sotto osservazione, nonostante la raccolta sia stata complessivamente stabile nel 2018. La maggiore attenzione è dovuta essenzialmente a due fattori:

- i) il più elevato premio al rischio sovrano che ha reso più costoso l'accesso al funding sul mercato;
- ii) la necessità di emissione di strumenti finanziari a medio/lungo termine volti al rispetto del requisito MREL per l'assorbimento delle perdite in caso di risoluzione o liquidazione.

Il contesto prospettico nell'anno in corso presenta alcuni fattori di rischio con potenziali impatti negativi sulla raccolta. L'aumento del premio sul rischio del debito sovrano, registrato nel 2018, ha mostrato effetti moderati sul costo della provvista bancaria grazie alla maggiore solidità mostrata dalle banche italiane rispetto alla media di quelle straniere.

Tuttavia, gli effetti in alcuni comparti, in particolare sul costo marginale della raccolta sul mercato istituzionale, si sono manifestati con una certa frequenza.

Nuovi incrementi dei premi per il rischio (come già verificati nel 2018 e nella prima parte del 2019 -grafico 2) genererebbero ulteriori effetti negativi sul bilancio delle banche.

Oltre alle perdite sul portafoglio titoli di Stato (circa il 46% dei titoli in portafoglio attualmente è classificato come "fair value through other comprehensive income" in cui ricadono i titoli disponibili alla vendita), un nuovo aumento del rischio sovrano si trasmetterebbe principalmente attraverso due canali.

Il primo, attraverso un ulteriore aumento del costo della raccolta sul mercato all'ingrosso generato sia dalla stretta correlazione tra credit spread del sovrano e degli emittenti corporate, sia dal legame esistente tra il rating dell'emittente sovrano e quelli delle banche residenti; il secondo, per l'avversione al rischio da parte degli investitori verso gli emittenti domestici.

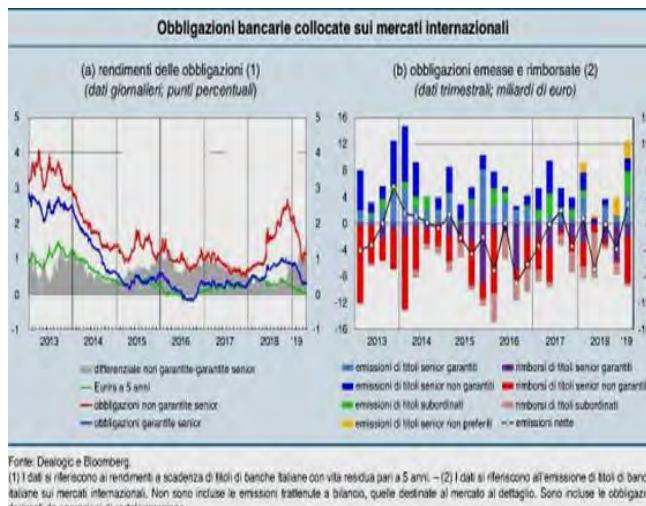
A ciò si aggiunge la perdita di valore del collaterale rappresentato da titoli governativi domestici che si può tradurre in una riduzione della capacità delle banche di rifinanziarsi presso l'Eurosistema o sul mercato monetario.

L'effetto complessivo è un peggioramento di accesso al funding come si è già verificato a partire dalla seconda metà del 2018 (grafico 3) con emissioni obbligazionarie, sui mercati internazionali da parte degli istituti di credito italiani, più costose.

Grafico 2 – Premio per il rischio sui titoli di Stato  
(Fonte Banca d’Italia)



Grafico 3 – Emissioni obbligazionarie sul mercato internazionale  
(Fonte Banca d’Italia)



## 1.2 Politica BCE a sostegno della crescita economica

In uno scenario di difficoltà economiche persistenti, la BCE ha lanciato tre programmi TLTRO: il primo, il 5 giugno 2014 (ripartito in otto aste), il secondo, il 16 marzo 2016 (ripartito in quattro aste) e il terzo, il 19 settembre 2019 (ripartito in sette aste).

L’autorità di Vigilanza Europea ha analizzato l’impatto esercitato nel tempo da questo tipo di interventi, evidenziando come i TLTRO rappresentino uno stimolo all’economia reale grazie alla stabilizzazione delle condizioni di finanziamento dal sistema bancario.

Inoltre, considerando le implicazioni che potrebbero verificarsi sul rispetto dei requisiti prudenziali di liquidità (le TLTRO sono una delle componenti del calcolo del Net Stable Funding Ratio), il nuovo rifinanziamento BCE fornisce una base importante per dare alle banche un supporto lato Funding con un contenimento dal lato costi, ma non risolve il problema di incrementare i ricavi in uno scenario di tassi negativi.

Il sistema bancario italiano, insieme a quello spagnolo, presenta attualmente la più alta incidenza del rifinanziamento BCE sul totale dell’attivo, pari rispettivamente a **6,5%** e **6,4%** (grafico 4).

Senza un nuovo TLTRO varato recentemente dalla Banca Centrale, a farne le spese sarebbero state soprattutto le banche italiane, che dovranno restituire 250<sup>2</sup> miliardi di euro nel 2020 (di cui 105 miliardi di euro concentrati presso le prime sette banche italiane - grafico 5).

Grafico 4 - Prestiti BCE alle Banche dell’Eurozona (Fonte BCE)

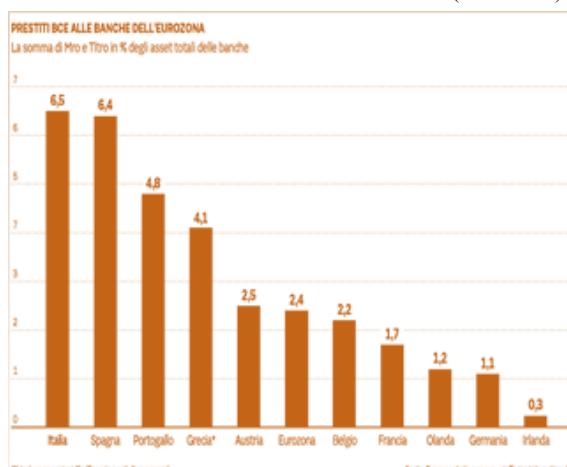
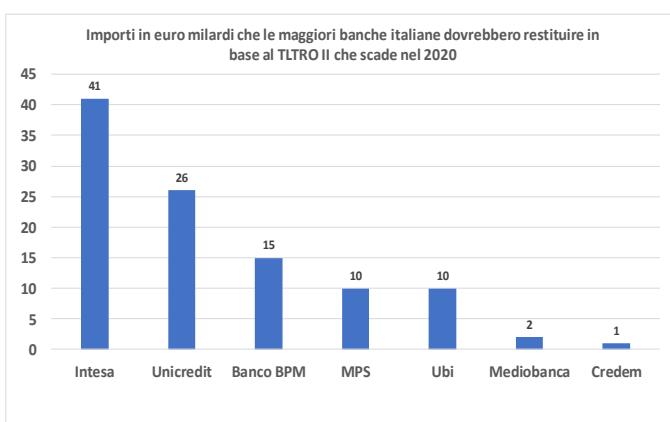


Grafico 5 - Prestiti BCE alle Maggiori banche italiane (Fonte BCE)



<sup>2</sup> l’importo da restituire ammonterebbe a 125 miliardi di euro, poiché la clausola *Net stable funding ratio* consente alle banche indebitate di ridurre del 50% l’importo da restituire nel 2020 mediante l’emissione di nuovi prestiti obbligazionari, purché tali operazioni siano state fatte entro il giugno 2019.

Le nuove operazioni **TLTRO III** disposte dalla Banca Centrale Europea in data **12.09.2019**, a condizioni più vantaggiose rispetto alla versione precedente (**TLTRO II**), **mirano a garantire condizioni al credito ultra-accomodanti**. Le sette aste di liquidità (da settembre 2019 a marzo 2021) comprendono prestiti di durata triennale e **il tasso applicato alle banche sarà uguale al tasso medio delle aste principali di riferimento**.

Ci sarà, inoltre, un tasso più basso per le banche che tra la fine del marzo 2019 e la fine del marzo 2021 supereranno il livello di prestiti di riferimento. Inoltre, sarà consentito alle controparti di rimborsare la somma presa in prestito prima della scadenza finale, con una frequenza trimestrale a partire da due anni dopo il settlement di ciascuna operazione.

A sostegno della crescita economica, l'Autorità di Vigilanza Europea ha anche deciso di tagliare il tasso di interesse sui depositi presso la Banca Centrale portandolo dal **-0,4% al -0,5%**. Al contrario, i saggi sulle operazioni di rifinanziamento principali e sulle operazioni di rifinanziamento marginale sono rimasti invariati rispettivamente allo **0,00%** e allo **0,25%**. Inoltre, dal 1° novembre partirà il programma di acquisto di titoli (APP) a un ritmo mensile di 20 miliardi di euro, senza una scadenza predefinita.

## 2 Crediti deteriorati e nuovi scenari alla luce delle nuove normative

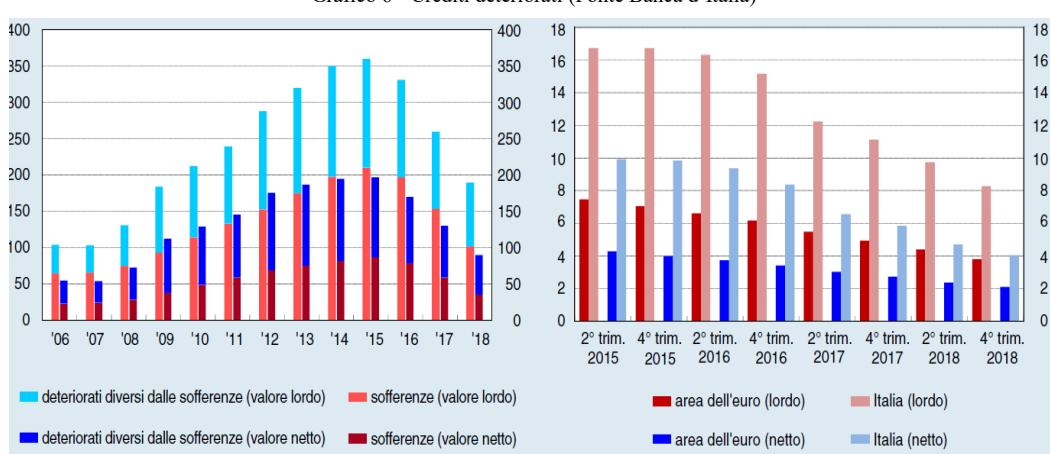
La crisi che ha colpito l'economia globale, a partire dal 2007, ha avuto importanti impatti sulle banche italiane, soprattutto per il deterioramento della qualità del credito.

Il tema della qualità del credito è stato oggetto di numerosi interventi regolamentari ma rappresenta, ancora oggi, uno dei problemi più complessi da risolvere.

Gli NPL, nonostante il calo significativo degli ultimi anni, sono una grande debolezza del sistema finanziario e determinano un deficit nella fornitura di credito all'economia del paese, con ricadute negative sulla ripresa economica.

Nel sistema bancario italiano, i crediti deteriorati, sebbene in riduzione dal 2016 (grafico 6), permangono su livelli particolarmente elevati e superiori rispetto all'area euro; per tale motivo la gestione delle esposizioni deteriorate è interesse primario delle banche italiane.

Grafico 6 - Crediti deteriorati (Fonte Banca d'Italia)



L'evoluzione del contesto macroeconomico, competitivo e regolamentare contribuirà a determinare il futuro andamento degli NPL. Al fine di rendere più solido il sistema finanziario, le autorità europee hanno varato alcune nuove norme/orientamenti sui crediti anomali di seguito sintetizzati:

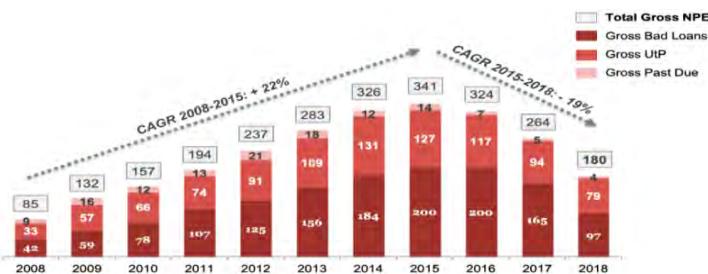
- la BCE ha pubblicato i nuovi criteri per la rilevazione dei crediti scaduti con arretrati di pagamento di 90 giorni consecutivi. Gli intermediari finanziari dovranno adottare la **"nuova definizione di default" entro il 1° gennaio 2021, ai sensi dell'art. 178 del Regolamento (UE) N. 575/2013**, anche ai fini della classificazione dei crediti come Non Performing Loan. La banca è tenuta a classificare un'esposizione in default quando il cliente è in arretrato da oltre 90 giorni su un'obbligazione creditizia rilevante. Per determinare la rilevanza dell'esposizione per le persone fisiche è stata identificata una soglia di rilevanza, articolata in due componenti: i) la componente assoluta pari a 100 euro<sup>3</sup> e ii) la componente relativa pari all'1% dell'importo totale delle esposizioni del cliente verso la banca finanziatrice. L'esposizione è classificata in default quando il suo arretrato supera, per un periodo superiore a 90 giorni, la soglia di rilevanza sia per quanto riguarda la componente assoluta che quella relativa. La banca sarà tenuta a classificare il cliente a default anche in presenza di disponibilità su altre linee di credito non utilizzate. Lo stato di default permanerà per almeno 90 giorni, dal momento in cui il cliente regolarizzerà l'arretrato di pagamento su mutui e/o prestiti e/o rientra dallo sconfinamento di conto corrente.

<sup>3</sup> Per le imprese, invece, la componente assoluta della soglia di rilevanza non è superiore a 500 euro.

- Il Parlamento Europeo, in data 14 marzo 2019, ha approvato le nuove norme UE per la copertura minima standard dei crediti in sofferenza (**Minimun Loss Coverage for Non Performing Exposures**). Le esposizioni non garantite dovranno essere coperte entro tre anni. Le esposizioni garantite dovranno essere coperte entro sette anni (se i collaterali sono strumenti del mercato mobiliare) o 9 anni (se le garanzie sono immobili).
- L'Autorità bancaria europea (EBA), in data 31 ottobre 2018 ha varato la versione finale degli **orientamenti sulla gestione delle esposizioni deteriorate e soggette a forbearance** (in vigore dal 30 giugno 2019). In particolare, gli orientamenti indirizzano le banche a concedere misure di forbearance con il solo obiettivo di far tornare il debitore a uno status in bonis e non utilizzare le ristrutturazioni del debito per ritardare il passaggio dello stesso tra quelli in sofferenza. **Inoltre, per la gestione degli NPL viene introdotta una soglia linda pari al 5%. Quest'ultima non indica un livello ottimale di NPL ma stabilisce un quadro prudenziale per salvaguardarsi da crescenti livelli di NPE.**

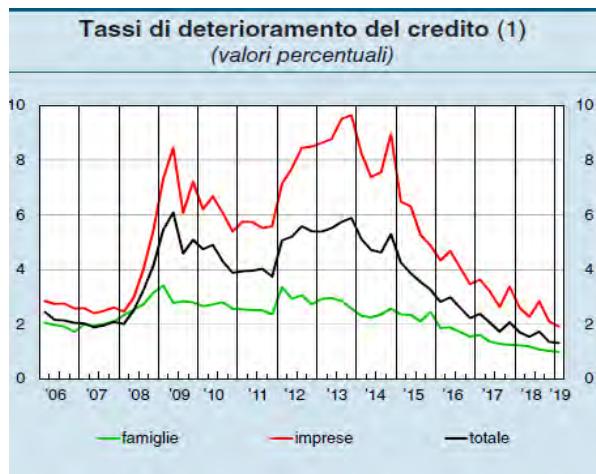
Nel 2018 i volumi dei crediti deteriorati nei bilanci delle banche italiane significative sono diminuiti per effetto delle rilevanti cessioni (grafico 7). L'attuazione dei piani di riduzione delle posizioni in sofferenza hanno contribuito all'ulteriore calo dell'incidenza del volume complessivo dei crediti deteriorati sul totale dei finanziamenti. Nel dicembre 2018, gli NPL delle banche italiane si sono attestati a €180 miliardi lordi (di cui 97€ mld di sofferenze, 79€ mld di UTP e 4€ mld di Past Due) rispetto ai €264 miliardi registrati nel dicembre 2017.

Grafico 7 -Evoluzione dello stock di Gross NPE (miliardi di Euro- Fonte Banca d'Italia)



Nel primo trimestre 2019 il flusso dei nuovi crediti deteriorati sul totale dei finanziamenti ha continuato a scendere rispetto al 2018 (grafico 8).

Grafico 8 – Tasso di deterioramento (Fonte banca d'Italia)



La BCE continuerà ad affrontare il tema delle consistenze degli crediti deteriorati e si confronterà con le banche interessate per definire aspettative di vigilanza a livello di singola banca, al fine di conseguire un livello di copertura coerente per le consistenze e per i flussi di NPL.

Gli istituti di credito potrebbero dover svalutare interamente i crediti anomali, e non solo quelli generati a partire dall'aprile 2018 come indicato nell'Addendum BCE (Addendum alle Linee guida della BCE per le banche sui crediti deteriorati). La Vigilanza Europea guarda, infatti, con occhio critico i crediti deteriorati raccomandando di implementare un graduale aumento dei livelli di copertura e svalutare integralmente lo stock entro la fine del 2026.

L'andamento futuro dello stock di Non Performing Loan nei bilanci degli istituti bancari sarà fortemente influenzato dall'effetto combinato di diversi fattori, quali:

- riduzione del tasso di decadimento e, quindi, dei flussi di nuove sofferenze;
- miglioramento della qualità dei processi di recupero;
- perfezionamento dei meccanismi di controllo;
- incremento delle cessioni di sofferenze sul mercato.

### **3 Nuovo requisito Mrel4: implicazioni per le banche alla luce dell'evoluzione della normativa BRRD2**

Il rafforzamento delle riserve di capitale e di liquidità dovrebbe ridurre la probabilità di future crisi e migliorare la resilienza degli enti in ipotesi di stress, siano esse causate da perturbazioni sistemiche o da eventi specifici riguardanti il singolo ente creditizio.

La Direttiva 2019/879/UE (BRRD II), che modifica la Direttiva 2014/59/UE (BRRD) e la Direttiva 98/26/CE, ha previsto per tutte le banche della UE un nuovo requisito MREL/TLAC al fine di assicurare che ogni intermediario finanziario, in caso di risoluzione, disponga di un ammontare di risorse patrimoniali e di altre passività in grado di assorbire le perdite e ricostituire il capitale.

La sua introduzione potrebbe dar luogo a un rilevante aumento del costo della raccolta per gli intermediari e a una minore disponibilità di credito bancario per l'economia, come segnalato nel rapporto sulla definizione e l'attuazione del requisito MREL pubblicato dall'Autorità Bancaria Europea (European Banking Authority, EBA) alla fine del 2016.

L'SRB (Single Resolution Board), nel 2018, ha suddiviso i piani di risoluzione in due cicli: il primo ciclo, concluso il 20 novembre 2018, ha riguardato le banche meno complesse mentre il secondo ciclo, concluso il 16 gennaio 2019, ha interessato le banche più complesse (istituti di credito con partecipazioni estere che implicano il coinvolgimento di collegi di risoluzione).

La policy 2018 di SRB ha definito sia il requisito MREL che il requisito minimo di strumenti subordinati. **Per le banche meno complesse** (cfr. 2018 SRB Policy for the first wave of resolution plans, del novembre 2018), identificate come altre istituzioni sistemiche (O-SII), la percentuale minima di strumenti subordinati dovrà essere pari al 12% dei RWA + il CBR (combined buffer) + NCWO add-on, mentre **per le banche più complesse** (2018 SRB Policy for the second wave of resolution plans), la percentuale minima di strumenti subordinati dovrà essere pari :

- i) al 16% dei RWA + CBR + NCWO<sup>5</sup> add-on, per le G-SII;
- ii) al 14% dei RWA + CBR+NCWO add-on per le altre istituzioni sistemiche (O-SII).

L'SRB determinerà obiettivi vincolanti caso per caso per le controllate di gruppi bancari complessi, dando la priorità alle entità più rilevanti (quelle che rappresentano più del 5% dell'esposizione a rischio del gruppo o del reddito operativo o che forniscono funzioni critiche all'intero gruppo).

Il 15 febbraio 2019 (dopo un iter legislativo durato circa due anni), il Coreper (Comitato dei Rappresentanti permanenti dei Governi degli Stati Membri dell'UE) ha adottato il testo finale della revisione della Direttiva BRRD.

Nel nuovo testo normativo BRRD2 (Direttiva 2019/879/UE del 20 maggio 2019) si evidenzia l'introduzione della categoria delle Top Tier Bank (banche con totale attivo superiore a 100 miliardi di euro), alle quali si applicherà un requisito MREL simile a quello delle G-SII. In Italia, in relazione agli ultimi dati di bilancio, sei gruppi bancari superano questa soglia, inclusa una G-SII e il nuovo gruppo cooperativo ICCREA<sup>6</sup>.

Il Parlamento Europeo, in data 16 aprile 2019 (Adoption of the banking package: revised rules on capital requirements (CRRII/CRD V) and resolution (BRRD/SRM)), ha approvato le modifiche alle Direttive CRR, CRD4, BRRD e SRM.

<sup>4</sup> TLAC – (Total Loss Absorbing Capacity), applicabile solo alle banche G-SII di tutto il mondo viene definito in Europa come Pillar 1 MREL. Queste banche devono disporre di determinati quantitativi di titoli in circolazione in grado di assorbire le perdite, che costituiscono nel complesso la TLAC delle singole banche. I nuovi requisiti in vigore nel 2019 dopo l'approvazione della BRRD2 e CRRII/CRD4 prevedono un livello TLAC pari al 16 per cento in termini di attività ponderate per il rischio (RWA) o al 6 per cento della leverage exposure. Entro il 1° gennaio 2022 queste quote saliranno al 18% delle RWA e al 6,75% della leverage exposure.

<sup>5</sup> Secondo tale principio, in caso di bail-in, nessun azionista o creditore della banca deve sopportare perdite maggiori di quelle che subirebbe in caso di liquidazione coatta amministrativa.

<sup>6</sup> Il Gruppo Bancario Cooperativo Iccrea nato il 4 marzo 2019 è formato da 142 Banche di Credito Cooperativo ed è il terzo gruppo bancario nazionale per sportelli e il quarto per attivi. Esso è stato costituito con la Legge n. 49 del 2016 (e successive modifiche), che ha riformato il Sistema del Credito Cooperativo e ha previsto l'obbligo di adesione delle Banche di Credito Cooperativo (tranne le Casse Rurali altoatesine) a una Capogruppo formalmente autorizzata dagli Organi di Vigilanza.

In seguito alle nuove disposizioni legislative, **entro il 28 dicembre 2020 dovranno essere recepite le modifiche alla BRRD** (Nuova Direttiva 2019/879/UE); nello stesso termine entreranno in vigore le modifiche al SRM. Fino ad allora, le decisioni prese da SRB si baseranno sulla politica MREL 2018 e sui requisiti SRMR1 / BRRD1.

Per le banche meno significative, il requisito MREL continuerà ad essere fissato dalle Autorità di Risoluzione Nazionali, in consultazione con le relative Autorità di vigilanza.

Alla luce della nuova BRRD (BRRD2), di seguito si sintetizzano i requisiti TLAC/MREL:

- **per le G-SII, ai fini del TLAC (in vigore dal 27 giugno 2019 - cfr. Minimum Requirement for Own Funds and Eligible Liabilities (MREL)),** viene richiesto un requisito MREL di primo pilastro in funzione del TREA e del LREM<sup>7</sup>. Le controllate di una G-SII, con sede fuori dall'UE e che non sono entità di risoluzione, saranno soggette ad un requisito di primo pilastro pari al 90% dei requisiti minimi previsti per la capogruppo con sede all'interno dell'Unione europea. E', inoltre, previsto un requisito MREL addizionale di secondo pilastro (in vigore dal 1° gennaio 2024 con un obiettivo intermedio entro il 1 gennaio 2022) specifico per singola banca;
- **per le Top Tier Bank,** così come previsto per le G-SII, il requisito MREL di primo pilastro richiesto sarà calcolato in funzione del TREA e del LREM<sup>8</sup> ma su percentuali **inferiori a quelli applicabili alle G-SII**. E' previsto un **requisito MREL specifico addizionale di secondo pilastro. I requisiti MREL di primo e secondo pilastro entreranno in vigore dal 1° gennaio 2024 con un obiettivo intermedio entro il 1° gennaio 2022;**
- **per le Other systemic entities (OSE),** le autorità di risoluzione avranno la facoltà di applicare il requisito MREL di primo pilastro in funzione del TREA e del LREM (nella stessa misura prevista **per le Top Tier Bank**) alle istituzioni bancarie che ritengano possano creare un rischio sistematico. Le singole OSEs sono anche soggette al requisito di secondo pilastro;
- **Per le altre istituzioni** (non G-SII, Top Tiers o OSEs), c'è invece solo un obbligo di requisito MREL di secondo pilastro. Le date di applicazione del requisito MREL (primo e secondo pilastro) delle **OSE e altre Istituzione Bancarie** (non G-SII, Top Tiers o OSEs) saranno in vigore anch'esse dal **1° gennaio 2024 con un obiettivo intermedio entro il 1° gennaio 2022.**

La BRRD2 (tab.2) prevede per le G-SII e le Top Tier Bank un livello benchmark di passività subordinate pari all'8% del TLOF (Total Liabilities and Own Funds), che può essere aggiustato in aumento o in diminuzione a discrezione delle autorità di risoluzione.

Questo benchmark è soggetto a un floor, pari al 13,5% dei RWA per le Top Tier Bank e al 14,5% per le G-SII. Inoltre, la BRRD2 prevede un cap del 27% dei RWA nel caso in cui la metrifica dell'8% del TLOF porti a un requisito di subordinazione superiore se misurato in termini di attivo ponderato per il rischio.

Tab. 2 – Requisiti di subordinazione (Fonte SRB e BRRD2)

Policy SRB 2018 2nd wave	BRD2
G-SII 16% RWA + CBR + NCWO add-on	G-SII 8% TLOF Floor 14,5% RWA + CBR Cap 27% RWA
Altre 14% RWA + CBR + NCWO add-on	Top Tier Bank 8% TLOF Floor 13,5% RWA + CBR Cap 27% RWA

Fonte: SRB, BRRD2

Per quanto indicato in tab. 2, si registra come la nuova BRRD2 fissa i livelli di subordinazione sul TLOF e non sul RWA. Le G-SII e le Top Tier Bank sono tenute a emettere passività subordinate per un valore pari ad almeno l'8% del TLOF, con la possibilità per le autorità di risoluzione, di accettare, in alcuni casi, livelli di subordinazione inferiori.

<sup>7</sup> Pari al 16% del Total Risk Exposure Amount (TREA) e al 6,00% del Leverage Ratio Exposure Measure (LREM) fino al 31 dicembre 2012 e del 18% del Total Risk Exposure Amount (TREA) e 6,75% del Leverage Ratio Exposure Measure (LREM) a partire da 1 gennaio 2022.

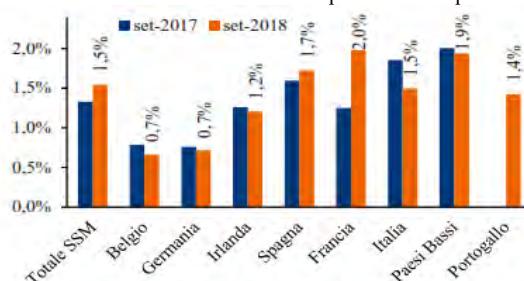
<sup>8</sup> Cfr. tab. 5 - pari al 13,5% del Total Risk Exposure Amount (TREA) e al 5,00% del Leverage Ratio Exposure Measure (LREM)

La logica per fissare i requisiti MREL subordinati con riferimento a una misura TLOF dell'8% è quella di migliorare la capacità di accedere ai fondi di risoluzione.

Rapportando i titoli subordinati al totale delle passività e dei mezzi propri, emerge che le banche italiane significative sono posizionate in linea con la media delle banche SSM, nel settembre 2018, ma la quota dei subordinati si è ridotta rispetto a dodici mesi prima (grafico 9).

Diversamente, le banche francesi e quelle dei Paesi Bassi mostrano una dotazione di subordinati superiore alla media.

Grafico 9- Peso % dei titoli subordinati sul totale delle passività e del patrimonio (Fonte dati BCE)



Fonte: nostre elaborazioni su dati BCE di 114 banche a settembre 2017 e 109 a settembre 2018.

**In riferimento a quanto indicato nel SRB 2nd way 2018, il requisito MREL per le banche significative è pari a: LAA + RCA + MCC=P1+P2R+CBR+P1+P2R+CBR-1,25% =2\* (P1+P2R+CBR) -1,25%**

- LAA: importo necessario per l'assorbimento delle perdite (Loss Absorption Amount, LAA= P1 + P2R + CBR) pari alla somma dei requisiti di capitale minimo richiesti dal primo pilastro (P1), dal secondo pilastro e al requisito combinato (Combined Buffer Requirement, CBR);
- RCA: importo necessario per implementare la strategia di risoluzione della banca (Recapitalisation Amount, RCA= P1 + P2R) pari alla somma dei requisiti di capitale minimo richiesti dal primo pilastro e dal secondo pilastro;
- MCC: importo necessario per assicurare la fiducia del mercato dopo la risoluzione (Market Confidence Charge, MCC. CBR-1,25%).

**Diversamente, per le banche non soggette a risoluzione il requisito MREL è pari a = LAA= P1+P2R+CBR**

Le autorità di risoluzione possono prevedere aggiustamenti, in aumento o in diminuzione, per RCA (includendo MCC) sulla base di un'approfondita analisi caso per caso, a seconda della situazione della banca e delle strategie di risoluzione individuate.

La BRRD2 ha previsto l'introduzione, per emittenti banche, di una nuova classe di obbligazioni chirografarie denominate Senior Non Preferred Bond. In Italia, questi titoli sono stati disciplinati con l'approvazione della Legge di Bilancio 2018, che ha introdotto le necessarie modifiche al Testo Unico Bancario.

Si tratta di titoli con durata pari a almeno un anno che non possono essere derivati né incorporare una componente derivata e che, nella gerarchia del passivo, sono di rango inferiore agli altri crediti chirografari e superiore alle obbligazioni subordinate.

Essi rappresentano un cuscinetto tra i bond subordinati Tier 2 e i bond senior preferred; la loro subordinazione rispetto a quest'ultimi si manifesterebbe solamente in una fase di risoluzione mentre non sarebbero convertiti o ridotti, come invece succederebbe alle obbligazioni subordinate, qualora venisse raggiunto il Point of Non-Viability (PONV).

Tale elemento fa ritenere che la probabilità di default dei titoli non-preferred senior sia identica a quella dei preferred senior, in quanto, entrambi gli strumenti entrano in default solamente in caso di risoluzione (rimane comunque maggiore la Loss Given Default per i titoli non-preferred).

Di seguito vengono riportate sia le passività ammissibili ai fini del calcolo del requisito MREL sia la gerarchia dei creditori:

- capitale di vigilanza (CET 1 e altre forme di capitale computabili come Tier1 e Tier2);
- titoli subordinati con durata residua maggiore di un anno;
- titoli Senior Non Preferred Bond con durata pari a almeno un anno;
- titoli obbligazioni senior non garantiti con durata residua superiore ad un anno;
- depositi con scadenza superiore ad un anno non idonei per la garanzia.

### 3.1. Impatto MREL sulle banche italiane

Per verificare gli effetti del MREL sul sistema bancario italiano, il CER (Centro Europa Ricerche) nel 2018 ha analizzato i bilanci (riferiti al 2017) delle prime nove banche in termini di totale attivo (Grafico 10- Fonte CER - Rapporto Banche 2-2018).

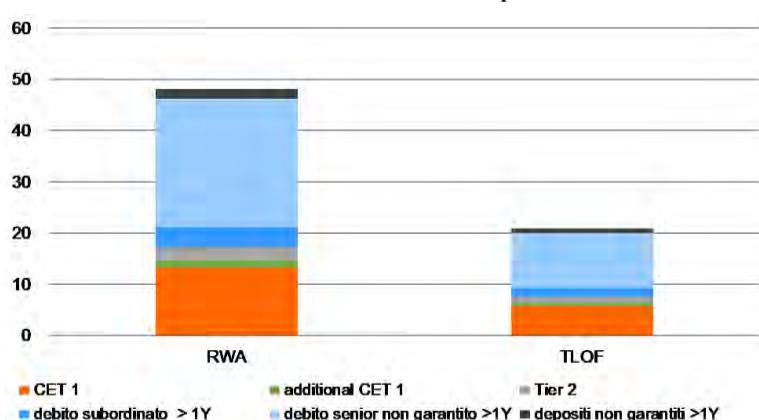
La composizione delle passività MREL rispetto alle RWA ha evidenziato, rispetto ai dati elaborati e stimati dall'EBA nel 2016, un valore pari al 48%, contro un livello medio per le stesse banche europee pari al 37,8%.

La percentuale di CET 1 capital in rapporto alle RWA è del 13%, mentre è pari al 25% per il debito senior non garantito, al 4% per il debito subordinato, al 3% per i depositi, 2% per il Tier 2 capital e 1% per l'additional CET 1 capital.

Il totale delle passività MREL, in rapporto al TLOF, è stimabile al 21%: valore più alto rispetto a quello medio per le banche europee individuato dall'EBA per il 2016 (14,4%).

In riferimento alle singole passività MREL, si nota come il debito senior non garantito con durata residua superiore ad un anno rappresenti la quota più ampia in percentuale del TLOF (11%), seguito dal CET 1 capital (6%).

Grafico 10- Passività MREL Banche italiane rispetto al RWA e TLOF



Le banche italiane sembrano avere una dotazione di passività MREL più elevata rispetto alle altre banche europee.

A seguito delle modifiche implementate sulla BRRD2, il MREL per le banche (G-SII e non G-SII) sarà espresso in percentuale degli RWA andando ad integrare i requisiti previsti da TLAC nelle norme generali già previste dall'Unione.

Al fine di poter verificare l'impatto del nuovo requisito MREL/TLAC, si riportano (tab. 3 e 4) i valori del Requisito MREL di alcune banche significative e del TLAC (applicabile solo a Unicredit in quanto definita banca G-SII). La stima dei suddetti valori è stata effettuata utilizzando i dati di bilancio del 2018\*.

Il requisito MREL evidenziato per alcune Banche significative (come definito nel SRB 2nd way 2018) è pari a:  $2*(P1+P2+CBR)-1,25\%$ .

TAB. 3- Requisito MREL di alcune Banche significative

BANCA	Pillar 1	Pillar 2 2019	TCR ( total Capital ratio)	CBR= Riserva di conservazione del capitale	a)= Requisito Mrel =2* (Pillar 1+Pillar 2 + CBR)-1,25%	Delta Requisito MREL Vs. TCR	b)= RWA in mln di euro* (IFRS 9 Transitional)	c)= FONDI PROPRI in mln di euro* (IFRS 9 Transitional)	d)= Passività eleggibili in mln di euro =(a*b)-c	e)= Mrel eligible instrument richiesto =a*b	f)= Totale passivo in mln di euro	g)= Requisito % TLOF =e/f
Gruppo Intesa	8%	1,50%	17,68%	2,50%	22,75%	5,1%	€ 276.446	€ 48.878	€ 14.013	€ 62.891	€ 787.721	8,0%
UBI	8%	2,25%	13,80%	2,50%	24,25%	10,5%	€ 61.035	€ 8.420	€ 6.381	€ 14.801	€ 125.306	11,8%
BPM	8%	2,25%	14,68%	2,50%	24,25%	9,6%	€ 64.324	€ 9.442	€ 6.157	€ 15.599	€ 160.339	9,7%

Dai dati di cui sopra, si registra per il Gruppo Intesa un requisito MREL più contenuto rispetto a UBI e BPM (22,75% Vs. 24,25%). Inoltre, si osserva come per UBI e BPM lo scostamento del requisito MREL sia significativo rispetto al TCR registrato nel 2018 (10,5%, e 9,6%).

Alla luce dei nuovi requisiti riportati nella tab. 9, si evince che gli istituti di credito con dimensioni minori (UBI, BPM) avrebbero maggiori impatti ai fini del rispetto del nuovo requisito MREL con probabili condizionamenti sia sui flussi che sul pricing dei crediti.

Nella tab. 4 viene riportato il requisito TLAC ( Pillar I MREL) per Banca Unicredit :

TAB. 4- Requisito TLAC Unicredit

BANCA	Pillar 1	Pillar 2 2019	TCR ( total Capital ratio)	CBR= Riserva di conservazione del capitale	a)= Requisito TLAC =16% +CBR	Delta Requisito MREL Vs. TCR	b)= RWA in mln di euro* (IFRS 9 Transitional)	c)= FONDI PROPRI in mln di euro* (IFRS 9 Transitional)	d)= Passività eleggibili in mln di euro =(a*b)-c	e)= Mrel eligible instrument richiesto =a*b	f)=Totale passivo in mln di euro	g) =Requisito % TLOF =e/f
UNICREDIT**	8%	2,00%	15,80%	2,50%	18,50%	2,7%	€ 370.180	€ 58.476	€ 10.007	€ 68.483	€ 831.469	8,24%

Il TLAC (in vigore dal 27.06.2019 – Addendum to the SRB 2018 MREL policy on new CRR requirements) è stato calcolato come un requisito transitorio pari al 16% + Combined Capital Buffer del totale degli RWA per raggiungere il requisito pieno del 18% + Combined Capital Buffer nel 2022 (senza alcun livello intermedio in termini di calibrazione).

Il nuovo requisito sopra indicato implica, con i dati di bilancio 2018, un ammontare di passività eleggibili pari a c.a. 10 mld di euro.

Al requisito TLAC di primo pilastro potrebbe aggiungersi un requisito addizionale di secondo pilastro in relazione alle caratteristiche specifiche dell’intermediario.

## Conclusioni

In un contesto economico che mostra segnali di affaticamento e con stime di crescita riviste al ribasso per i paesi dell’area euro (e non solo), le banche devono fare i conti con alcune regole adottate di recente in Europa e altre che entreranno in vigore, tutte mirate a rafforzare il sistema bancario e ridurre i rischi per l’economia ma che in realtà sono controproducenti. In prospettiva queste modifiche normative renderanno le banche meno rischiose, dotandole di più capitale, di maggiore liquidità, di un grado di leva più basso che in passato. Allo stesso tempo, le nuove regole comprimono la redditività bancaria e lo sviluppo del sistema creditizio; ne deriveranno riflessi negativi sulla disponibilità di prestiti all’economia reale. I cambiamenti avvengono, peraltro, in un contesto europeo di persistente debolezza economica e di basso autofinanziamento delle imprese, in presenza di un ruolo limitato dei mercati. Nei prossimi anni, l’introduzione del requisito MREL potrebbe dar luogo a un rilevante aumento del costo della raccolta per gli intermediari finanziari e a una minore disponibilità di credito bancario per l’economia, come documentato nel rapporto sulla definizione e l’attuazione del requisito MREL pubblicato dall’Autorità bancaria europea (European Banking Authority, EBA) alla fine del 2016. Inoltre, a seguito dell’entrata in vigore delle nuove norme sui crediti anomali, si prospettano maggiori coperture e di riflesso più cessioni di prestiti, in particolare per quelli non garantiti: segmento, questo, su cui potrebbe registrarsi un cambio di strategie da parte delle banche, con un significativo rialzo del tasso nei confronti dei prenditori (in particolare per le imprese). In questo scenario potrebbe mutare anche il business model degli istituti. Alla luce del nuovo contesto normativo, il recupero di redditività passerà per una serie di scelte strategiche e operative, determinate anche dall’azione di vigilanza:

- Diversity in diversification: accelerazione verso attività a minor assorbimento di capitale;
- Efficiency: recupero di efficienza economica e di crescita dimensionale centrale anche per ammortizzare i crescenti costi tecnologici e di compliance regolamentare;
- De-risking,: gli impatti principali sull’economia reale sono collegati alla riduzione delle RWA e di NPE.

Camillo Giliberto

## Bibliografia

Direttiva 2019/879/UE (Nuova BRRD2);  
 Direttiva 2014/59/UE (BRRD);  
 Direttiva 98/26/CE  
 2018 SRB Policy for the first wave of resolution plans;  
 2018 SRB Policy for the second wave of resolution plans;  
 CRR Addendum to the 2018 SRB Policy for the second wave of resolution plans  
 Minimun Loss Coverage for Non Performing Exposures (European Parliament)

# Classification of NPL with a Random Forest approach

di Massimiliano Zanoni (Iason Ltd)

Articolo sottoposto a doppio referaggio anonimo, pervenuto in data 23/12/2019 e accettato il 17/03/2020

## Abstract

Artificial Intelligence has quickly entered in the financial services industry covering a wide range of applications. This work studies a structured statistical approach to classify non-performing unsecured commercial exposures according to their recovery potential, based on a Machine Learning technique known as Radom Forest.

The framework adopted is based on two different components: one identifying the cases that may be recovered and the other estimating their recovery level. In addition, the work compares the RF - introduced with a review of the underlining Decision Tree theory and its performance metrics - to the better-known Logit approach.

The framework is meant to provide an evaluation of recovery at aggregate level, for pricing and management purposes, but it is also successfully tested in comparison between two portfolios, one of which know to the analyst. Results show that the Random Forest approach is as reliable and slightly more performing than the better known Logistic approach, even with a limited set of information.

## 1 Introduction

Learning is a process meant to identify patterns and matching rules within available data, in order to infer the inputs-outputs mapping. Traditional modelling performs this task by presuming a functional form between the two, then estimate the key parameters through a procedure called fitting.

**Statistical learning** instead consists in a set of methodologies where no functional form is assumed and classification is based directly on the characteristics of events (Gareth, et al., 2013). **Machine Learning** (ML) constitutes a set of non-parametric, non-linear statistical methods to implement statistical learning.

In the case considered, the training dataset consists of short-term non-performing commercial exposures to be recovered by a servicer, which usually purchases such portfolios at a discount and then works them out, hence a correct price estimate and an efficient workout process are key to profitability.

It is worth noticing that, though forecasting occurs at single deal level, what really matters to the investor is the ability to correctly infer the amounts which can be recovered at aggregate levels.

The underlying assumption is that both the recovery event and the recovery rate can be inferred from past transactions and debtor information. In particular, the work confirms, with (Khieu, et al., 2012), that loan's characteristics are more significant determinants of the recovery rate than are borrower characteristics prior to default. On the other hand, we were not able to inquire if predictors referring to the recovery process of the processor, or the bank prior to portfolio sale, were relevant as suggested by (Bellotti, et al., 2019), since such information was not available.

The approach adopted mirrors the standard Loss Given Default (LGD) framework which splits the default event from loss estimates and was designed independently of existing literature. Focusing on the servicer, the work complements the literature which mostly estimates recovery models from the same bank originating the loan (Ciavoliello, et al., 2016). In addition, it contributes to update the benchmark study involving Machine Learning methods.

Besides its good predictive capabilities, RF can provide valuable insight into the main factors driving the recovery dynamic (Breiman, 2001). It is worth noticing that, though forecasting occurs at single deal level, what really matters to the investor is the ability to correctly infer the amounts which can be recovered at aggregate levels (e.g. at geographical area level).

After an introduction to the RF approach, a chapter is dedicated to the description of the Non-Performing Loan (NPL) dataset before analysing the modelling approach where the model for the probability of recovery is compared to a classical generalized Logistic model.

Finally, the last chapter presents the results at regional and portfolio level, showing that the RF framework is as reliable as the better-known Logistic approach.

## 2 Tree-based clustering

Random Forest is a machine learning approach which sets binary rules, represented as trees, to cluster the event space by recursive branching. A clustering rule is a set of splitting points in each dimension of the event space, dividing (branching) the dataset into two separate parts. A set of rules represent a tree ending in a defined boxing of the event space (terminal node or *leave*) to which a unique value is assigned (prediction). All the different trees together are called the *Forest*. The final model is obtained by averaging the value assigned to any event by over all the splitting trees composing the forest.

While the linear approach relies on a 'linearized' law, the RF does not need any a-priory knowledge of the system. For each observation, a prediction is obtained by recursively applying the following steps:

1. Randomly select  $m$  variables from all the ones in the dataset (*feature bagging*).
2. Randomize the selection of variables in each split<sup>1</sup>

<sup>1</sup> The continuous change of variables at each split limits the chance to include correlation in the learning process.

3. For each of the selected variables identify the split-point that minimizes the residual sum of squares of predictions in the specific tree. Formally, for a randomly selected feature  $j$ , two regions are defined by the split point  $c$

$$(1) \quad R_1(j, c) = \{X | X_j < c\} \text{ and } R_2(j, c) = \{X | X_j \geq c\}$$

such that:

$$(2) \quad \sum_{i:x_i \in R_1(j,c)} (y_i - \hat{y}_{R1})^2 + \sum_{i:x_i \in R_2(j,c)} (y_i - \hat{y}_{R2})^2$$

is minimized, where  $\hat{y}_{R1}$  and  $\hat{y}_{R2}$  are average estimated outputs for each split.

4. Take the arithmetic average each outcome over all trees in the forest.

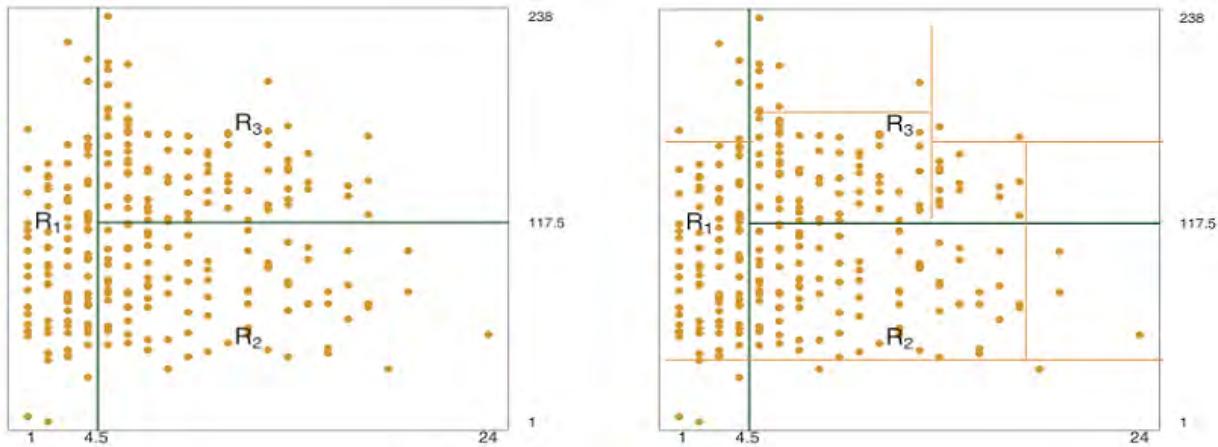


Figure 1 Splitting Process

The process is affected both by the number of variables tried at each split, and by the number of processed layers and it eventually ends up drawing a number of boxes, identified by one value of the variable, independently of the number of events included.

This process differs from that of bagged trees where all predictors are considered at each split, creating highly correlated tree. RF process overcomes this problem by randomly selecting a limited number of predictors at each split, which makes the model less prone to overfitting problem (Hastie, et al., 2009).

## 2.1 Performance indicators

Performance measures are key to drive variable selection and model validation. Beside classical statistical measures such as

- **Percentage Variance explained** (PVE): it measures how close the predictions on new observations get to the real variable. Formally:

$$(3) \quad PVE = \frac{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n} - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}$$

where  $\bar{y}$  and  $\hat{y}_i$  are the average of real and estimated outputs respectively<sup>2</sup>.

- **Receiver operating characteristic** (ROC)<sup>3</sup> curve (Figure 2), which is a usual performance measure for binary models. Formally Specificity or true negative and Sensitivity or true positive are defined as:

$$(4) \quad \text{Specificity} = \frac{A}{A+B} : P(Q = 0 | Y = 0), \text{Sensitivity} = \frac{D}{C+D} : P(Q = 1 | Y = 1)$$

In general, the recognition of events in model application requires the definition of a cut-off value that separates positive from negative cases. Together with their ‘false’ classification, they can be represented through a Confusion Table represented in Table 1.

<sup>2</sup> For a more rigorous disruption of variable importance in RF models refer to (Grömping, 2009) and (Breiman, 1996)

<sup>3</sup> For detailed analysis of the ROC method refer to (Fawcett, 2006)

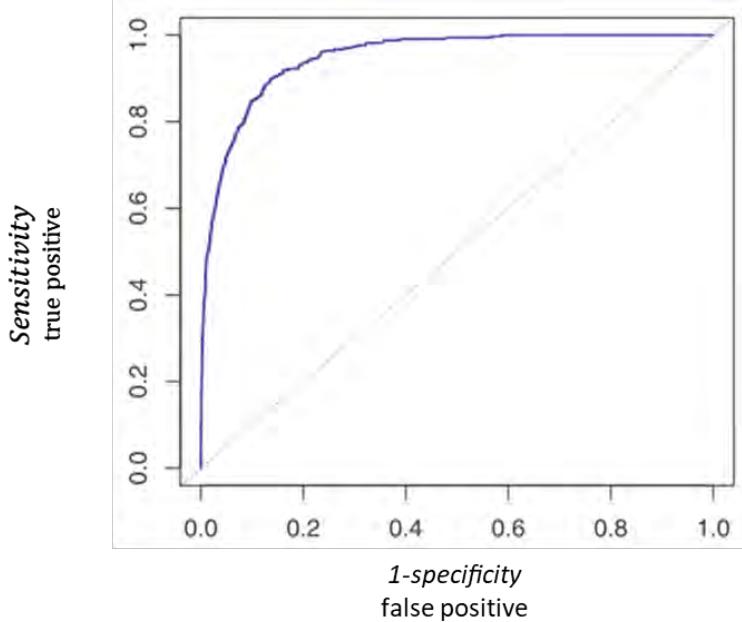


Figure 2 ROC Curve

		Predicted	
		Negative (0)	Positive (1)
Observed	Negative (0)	A	C
	Positive (1)	B	D

Table 1 Confusion Table

- **Mean decrease in accuracy** and **Mean decrease in node impurity** represent important measures of the contribution of each single variable; the former analysing the change in the prediction error due to the change in values of a given variable and the latter by comparing the “purity of the node split” in the alternative of including or excluding the given variable from the set used for splitting<sup>4</sup>.

For completeness, the Confusion Table is extended to include the totals of each measures, the portfolio mix, real forecasted and the diagonal sum of true positive and negative.

		Fit		Total	Portfolio-mix
		Non-Rec	Recovery		
Real	Non-Rec	20,523	354	20,877	97.0%
	Recovery	524	123	647	3.0%
		Total	21,047	477	21,524
		Portfolio-mix	97.8%	2.2%	
		OK	97.5%	25.8%	
		KO	2.5%	74.2%	Tot. acc.
		OK on sample	98.3%	19.0%	95.9%

Table 2 Extended Confusion Table – Example Linear Cut Off 10%

In the example (Table 2), the portfolio mix offers a rough measure of the model’s ability to replicate the original portfolio, forecasting only 2.2% of recovery events versus an effective 3%. Furthermore, it underlines the fraction of correctly separately forecasted success and fail cases (OK) and overall (OK on sample)<sup>5</sup>.

<sup>4</sup> For detailed definition refer to (Breiman, 2001)

<sup>5</sup> The overall number however does not distinguish between success and fail events, distinction that may be vital in cases where the cost of different mistakes is not homogeneous.

### 3 Dataset description

The NPL dataset includes 22,290 non-performing retail exposures related to unpaid utility bills, with limited information about the customer and the exposure provided both as snapshot at a given date and as payment flows. Some key information is inferred from the fiscal code (gender and age), hence exposures where this is missing are discarded. Several data quality issues imposed more exclusions, e.g. exposures related to debtors above 95.

As summarized in the following figure, recovery rates are distributed mostly towards extreme values, i.e. most dossiers present no recovery or a high recovery rate, leaving only a fraction in the middle of the recovery scale. In order to reduce statistical noise dossiers with a very small recovery rate or absolute amount, were also excluded from the analysis.

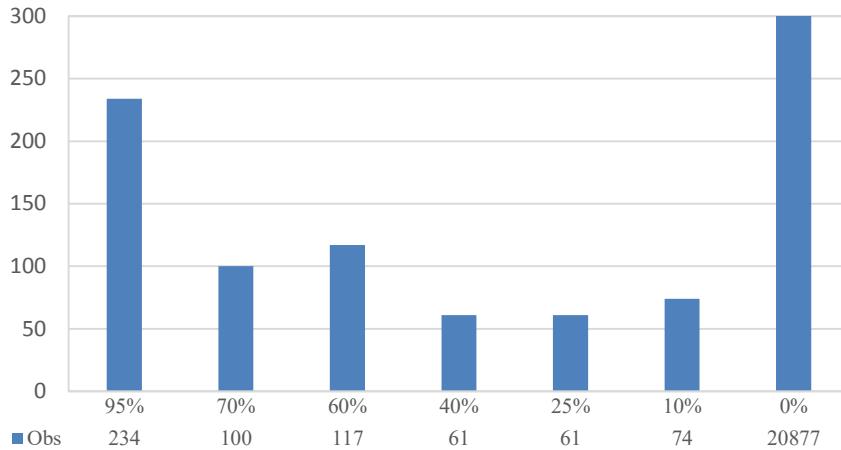


Figure 3 Number of observations by Recovery Level

The variable “*contacts*” (shown in the following histogram) represents the number of interactions with the counterparty in the recovery process, intuitively this is linked to the duration of the process and to the average recovery rate, at least up to a certain number of contacts. On average, dossiers with 140-170 contacts present a 50-55% recovery rate however, increasing the number of contacts does not improve the rate of recovery (in fact it is reduced to around 30-35%), showing a possible inefficiency when a significant effort is made on cases yielding a poor and slower recovery.

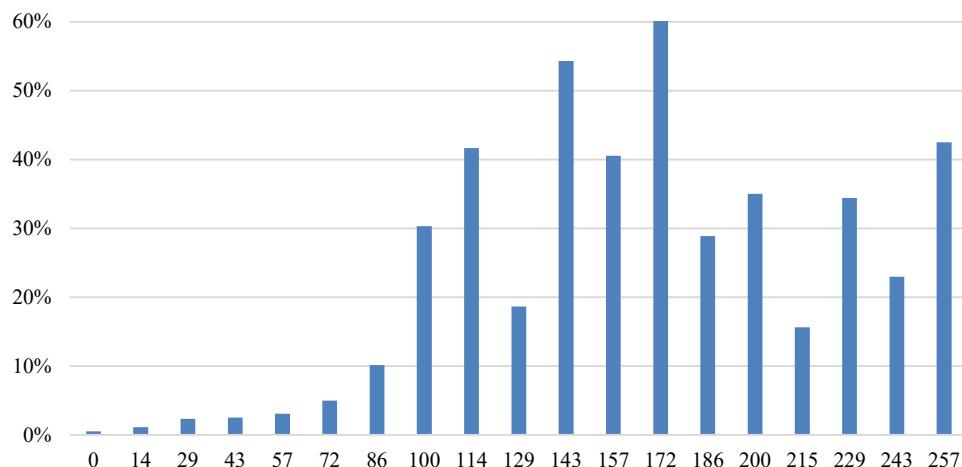


Figure 4 Average Recovery Rate by Number of Contacts

In many cases the recovery process involves a single flow but, in general, the cumulated recovery increases with time up to the final recovery amount. At any given period  $T$ , the *average cumulated recovery rate* is given by the sum of payments, divided by the total exposure of dossiers.

$$(5) \quad \text{average cumulated recovery rate} = \frac{\sum_{t,T} \text{payment}_t}{\sum_T \text{exposure}_T}$$

Given the different recovery dynamics, the analysis of the average recovery curve at portfolio level should focus on dossiers with homogeneous durations i.e. by restricting the recovery process to dossiers with a recovery duration above  $D$ ,

$$(6) \quad \frac{\sum_{t,T,K>D} \text{payment}_{t,K}}{\sum_{T,>D} \text{exposure}_{T,K}}$$

where  $K$  selects those dossiers with a collection process exceeding duration  $D$ .

Clearly with D=0 the set embraces the whole portfolio, including all the quick-recovery dossiers, with each set reflecting an average of many dossiers with different recovery dynamics and durations.

The curve is steeper in the initial part, but reaches a lower total recovery rate (blue line in Figure 5) with respect to subsets that includes dossiers with a longer recovery process, e.g. over 3 months (orange line) which presents a slower rise to a higher recovery rate.

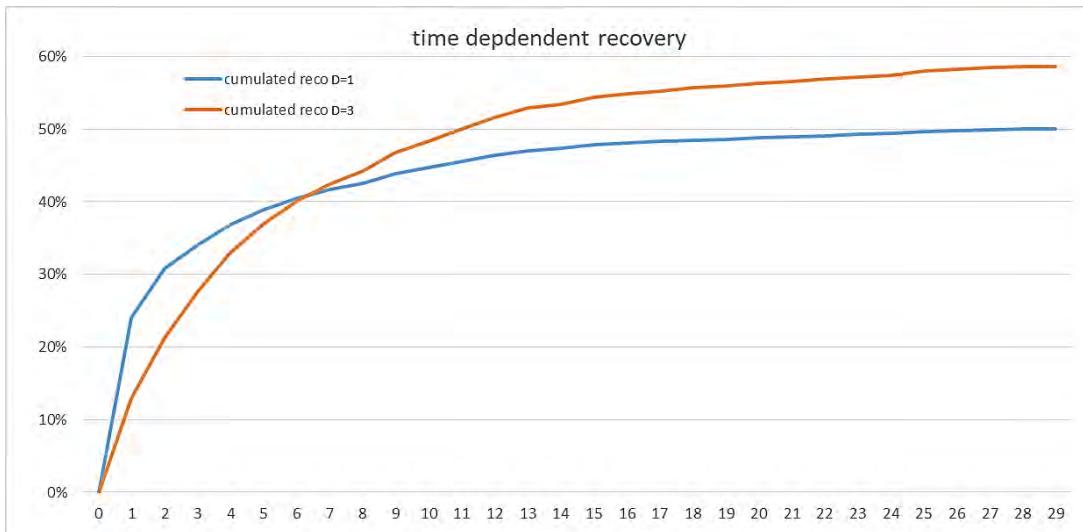


Figure 5 Time Dependent Recovery: Average Cumulated Recovery Rate Dynamic for different portions of portfolio

## 4 Development approach

The framework used to forecast recoveries in a portfolio of NPL exposures is based on two different models: one dedicated to forecast the probability of recovery of the single dossier and one to estimate the recovery rate in case of a recovery. Accordingly, the estimation process is divided into two parts: the first one is meant to estimate the *Probability of Recovery* (PR) of a given dossier; while the second is targeted to estimate the Recovery Rate of each recovered exposure.

The two models are estimated on the same variable set and tested separately, they will thus be statistically independent of each other, but dependent on the availability and quality of the same dataset.

In order to provide a frame for comparison, beside the RF approach, the PR model is estimated also using a generalized linear approach (Logit) while, for the recovery rate, only one multi-linear model is estimated.

Each of the two models is characterized by a specific performance level. Thus it will introduce an independent error into the final recovery estimates.

However, since the two models are statistically independent, their errors may partially cancel out when dossiers are aggregated at macro-regional level, potentially generating a more accurate total recovery with respect to what obtained at dossier level.

The performance measure used for the Recovery given Recovery (RGR) model will be a classic adjusted R-squared correlation, while to measure the quality of the PR model and compare the two approaches, a ROC curve, together with an extended Confusion table, is used, in addition to specific metrics evaluating the relevance of each variable (Percentage of Variance Explained and Increased Node Purity).

## 5 Probability of recovery model

The probability of recovery as a continuous variable, is estimated similarly within the Logit and the Random Forest approach, then transformed in a list of recovery events through a specific cut-off optimized for each approach. Alternatively, recovery events can be tagged directly by using the RF as a classifier. This result will be compared to the one obtained from the continuous PR.

### 5.1 Threshold definition

The cut-off necessary to determine which dossier should be considered a successful recovery can be defined in different ways:

- A. A possible value is obtained by dividing the dataset in real recovery and non-recovery events and plotting each group along the probability of recovery assigned by the model. The value at the intersection point of these two distributions can be taken as cut-off (Figure 6).

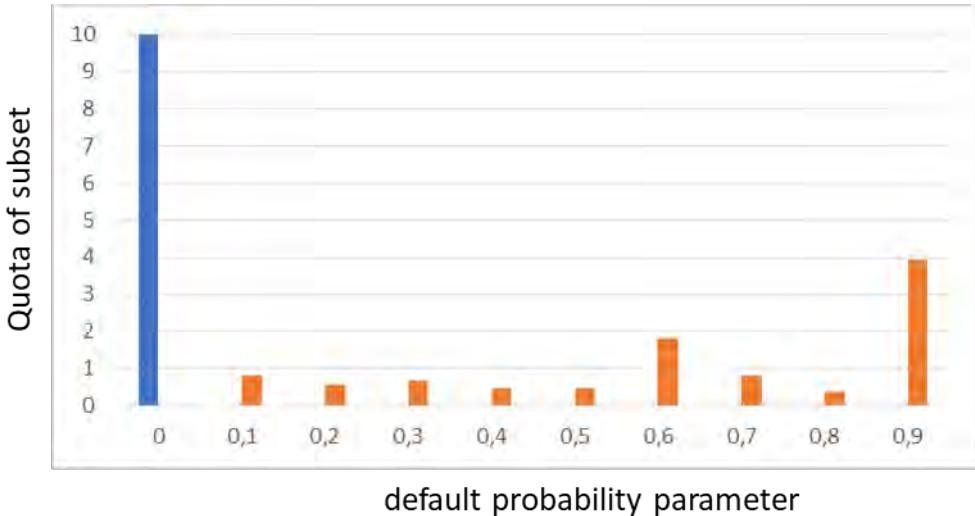


Figure 6 Quantile Distribution of Recovery (Orange) and Non-Recovery Evens (Blue)

B. An alternative and more rigorous way to define the cut-off is to set it in a way that the number of estimated recoveries is close to the real one.

C. Table below shows the number of predicted recoveries as a function of the cut-off in the portfolio analysed including 626 recoveries. This approach suggests the same cut-off selected with the previous visual approach i.e. PR = 0.1.

cut-off >=	predicted
9%	831
10%	655
11%	512

Table 3 CUT-OFFS

D. An alternative analytic approach is to define the cut-off that maximizes the overall number of success cases - both negative and positive – i.e. the value that alternatively:

- minimizes the differences (Specificity – Sensitivity)<sup>6</sup> considering positive values only
- maximizes the sum (Specificity – Sensitivity) commonly known as Youden's Index (Youden, 1950)

Considered that in the present work, method **A** and **B** identify the same cut-off level, only the two cut-offs will be compared.

## 5.2 Linear approach

The model estimated with the Logit approach is summarized in the following table:

Variable	P-value	relevance
Intercept	0.000	High
Abroad	0.000	High
Debtor age	0.000	High
Macro Region: Center-South	0.005	Medium
Macro Region: Island	0.037	Low
Macro Region: North	0.141	Low
Contact	0.000	High

Table 4 Logistic Model for the Probability of Default

<sup>6</sup> See equation (4) for relevant definition

The ROC curve associated to this model visually shows how much it improves the random guess represented by the straight line. Though the ROC does not depend on a specific cut-off, the performance of the model, in terms of correctly identified recovery events, is based on a cut-off chosen.

In more details, the performances of the different models and options are compared through an extended Confusion matrix which includes, besides true/false positive and negative events, the following metrics:

- **Total accuracy** (i.e. correct predictions) -  $(TN + TP) / \sum \text{all}$
- **Precision** =  $TP / (TP + FP)$  - (# of predictions)
- **Specificity** =  $TN / (TN + FP)$  (# num of negative occurrences)

The Confusion matrices in Table 5 and Table 6 allow to compare option **B** and **C** outlined before<sup>7</sup>. Adopting definition **B** for the cut-off (10%) the resulting portfolio mix is very similar to the real one (3.1% vs 2.9%) but only 21.5% of recovery signals is correct (Table 6).

Adopting definition **C** for the cut-off (3%), which grants the same performance for both positive and negative forecasts, a much larger set of recovery events is identified, at the expenses of a significant increase in wrong recovery signals (94.1%, vs 78.5% of portfolio), leading to an excessive 33.7% portfolio mix.

At the base of this approach, is not only the wrong assumption that the two event types cover, more-or-less, the same portion of the population, but also both that the same cost is associated to both types of mistakes, clearly incorrect.

In summary, approach **B**, identifying a number of recoveries close to the real one (655 vs 626), is used to compare the RF to the Logit approach, as it provides a good estimate of the portfolio mix.

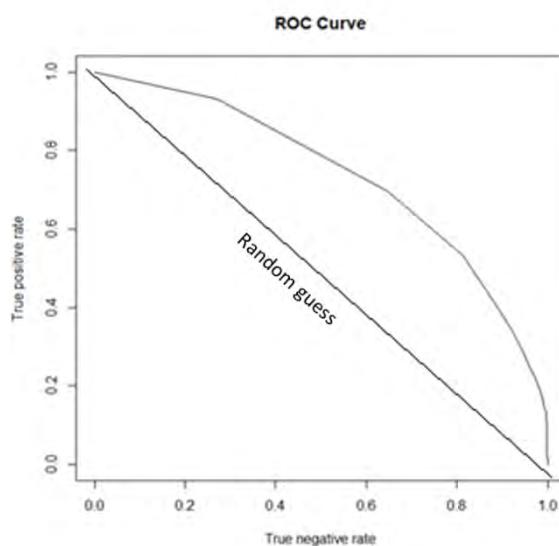


Figure 7 ROC Curve – Logistic Approach

		Fit		Portfolio-mix
		Non-Rec	Recovery	
Real	Non-Rec	20,094	514	20,608
	Recovery	485	141	626
	Total	20,579	655	21,234
	Portfolio-mix	96.9%	3.1%	
	Precision	97.6%	21.5%	
	KO	2.4%	78.5%	Tot. acc.
		97.5%	22.5%	95.3%

Table 5 Linear approach, cut-off method B

<sup>7</sup> Option A is not reported as it resembles option B.

		Fit		<b>Total</b>	Portfolio-mix
		Non-Rec	Recovery		
<b>Real</b>	Non-Rec	13,879	6729	20,608	97.1%
	Recovery	205	421	626	2.9%
	<b>Total</b>	14,084	7150	<b>21,234</b>	
	Portfolio-mix	66.3%	33.7%		
	<b>Precision</b>	98.5%	5.9%		
	<b>KO</b>	1.5%	94.1%	<b>Tot. acc.</b>	
<b>OK on sample</b>		67.3%	67.3%	<b>67.3%</b>	

Table 6 Linear approach, cut-off method C

## 6 Random Forest approach

A different model for the probability of recovery was estimated on the same dataset using a Random Forest approach in regression mode. The output is then turned into recovery events with a new cut-off, calibrated on the same principle, i.e. to obtain a number of recovery events closer to the real one.

		Fit		<b>Total</b>	Portfolio-mix
		Non-Rec	Recovery		
<b>Real</b>	Non-Rec	20,132	476	20,608	97.1%
	Recovery	468	158	626	2.9%
	<b>Total</b>	20,600	634	<b>21,234</b>	
	Portfolio-mix	97.0%	3.0%		
	<b>Precision</b>	97.7%	24.9%		
	<b>KO</b>	2.3%	75.1%	<b>Tot. acc.</b>	
<b>OK on sample</b>		97.7%	25.2%	<b>95.6%</b>	

Table 7 RF approach, cut-off method B

As shown in Table 7, the RF approach identifies a higher number of true recovery cases than the Logit approach (158 vs 141) in a smaller set classified as recovered, thus hitting two scores: a portfolio mix closer to the real one, and a higher performance (24.9% vs 21,5%). It is worth noticing that, differently from the Logit approach, the RF approach includes a non-deterministic component that changes the output at each new run, generating a slightly different classification. With the RF approach, the model usually includes all variables with different relevance, as shown in the following table, where they are ranked according to their relative importance<sup>8</sup>.

Variable	IncNodePurity	%IncMSE
Exposure	20,23	2,18E-03
Contacts	48,33	7,60E-03
Dossier age	11,87	6,84E-04
Macro region	2,45	2,86E-04
Abroad/Italy	1,00	6,60E-04
Debtor age	8,07	5,88E-04
Debtor gender	0,70	1,85E-05

Table 8 Variable performance – RF-PR model

<sup>8</sup> The performance and robustness depend on parameters which need to be tuned properly during the estimation process, the ones used in the present model are: the number of trees, (500) minimum size of terminal nodes in each tree, (50) the number of variables to be selected every time, (2)

As shown, the exposure is the most relevant variable in the model, fact that may create a preference for larger exposure, with the possible result of overestimating recovery exposure. The Logit and the RF models can be directly compared through their relative ROC curves, both plotted on the same graph.

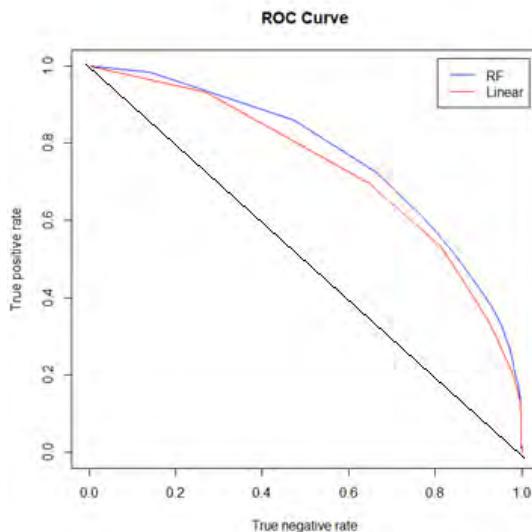


Table 9 ROC curve – comparison of RF & Logistic approach

## 7 Random Forest as direct classifier

Beside “regression mode” the RF approach can be used as a classifier, with no need to define a cut-off. Recovery events are directly determined by the branching algorithm. The outcome is summarized in Table 10.

		Fit		<b>Total</b>	Portfolio-mix
		Non-Rec	Recovery		
<b>Real</b>	Non-Rec	20,567	41	20,608	97.1%
	Recovery	541	85	626	2.9%
<b>Total</b>		21,108	126	<b>21,234</b>	
Portfolio-mix		99.4%	0.6%		
<b>Precision</b>		97.4%	67.5%		
<b>KO</b>		2.6%	32.5%	<b>Tot. acc.</b>	
<b>OK on sample</b>		99.8%	13.6%	<b>97.3%</b>	

Table 10 RF approach as a classifier

This approach produces the best results by far in terms of true forecasting performance (67.5% vs the usual 22-24%). However, this comes at the expenses of an extremely careful classification of positive events: only 126 are selected, generating a very different portfolio mix from the real one (0.6% vs 2.9%). Consistently, the large number of undetected recoveries reduces the performance on the recovery set 13.6% (vs 25.2% obtained in regression model).

## 8 Overfitting test

Overfitting happens when a model learns the noise in the data to the extent that the performance of the model on new data is poor. To test this feature, the given portfolio is divided into two sub-portfolios, referred to as train and test portfolios. The characteristics of the two portfolios are summarized in Table 11. A new model is estimated on the train portfolio with the same approach. To keep the number of recoveries close to the actual number of recovered dossiers, the 12% cut-off is chosen, yielding 415 fitted recoveries.

	Dossier	Recovery	Prob. Recov	Exposure/000	Recovery/000
<b>Train Set</b>	14,014	413	2.95%	14,907	284
<b>Test Set</b>	7,220	213	2.95%	7,993	152
<b>Total</b>	21,234	626	2.95%	22,900	437

Table 11 Test and Train Portfolios

		Fit			
		Non-rec	Recovery	Tot	Portfolio-mix
Real	Non-rec	13,285	316	13,601	97.1%
	recovery	314	99	413	2.9%
	Total	13,599	415	<b>14,014</b>	
	Portfolio-mix	97.0%	3.0%		
	Precision	97.7%	23.9%		
	KO	2.3%	76.1%	Tot. acc.	
OK on sample		97.7%	24.0%	<b>95.5%</b>	

Table 12 Extended Confusion Matrix for train Portfolio

		Fit			
		Non-rec	Recovery	Tot	Portfolio-mix
Real	Non-rec	6,825	182	7,007	97.0%
	recovery	153	60	213	3.0%
	Total	6,978	242	<b>7,220</b>	
	Portfolio-mix	96.6%	3.4%		
	Precision	97.8%	24.8%		
	KO	2.2%	75.2%	Tot. acc.	
OK on sample		97.4%	28.2%	<b>95.4%</b>	

Table 13 Extended Confusion Matrix for Test Portfolio

Table 12 and Table 13 present the extended confusion matrix for train and test portfolios. The test portfolio offers a rough measure of the model's ability to replicate the train portfolio, forecasting 3.4% of recovery events versus 3%. Furthermore, it underlines approximately similar fraction of correctly separately forecasted success and fail cases (OK) and overall (OK on sample).

## 9 Recovery rate model

The model estimating the recovery rate, for those dossiers for which a recovery occurs, is developed with a standard multilinear approach, on the basis of information available on relevant recovered dossiers. For this model, no RF alternative is provided. The recovery rate (RR) is defined in the same way at granular and aggregate level as:

$$(7) \quad RR = \frac{\text{Recovered Amount}}{\text{Exposure}}$$

The geographic variables (macro region) are maintained in the model summarized below, even with low relevance, as they characterize the final level of aggregation. The *Adjusted R-squared* (10.4%)<sup>9</sup> of the selected model confirms that the overall performance is quite low but this does not prevent a good forecast of overall recoveries at aggregate level (macro-region), as shown in Table.

Variable	P-value	Relevance
Intercept	0.000	High
Exposure	0.000	High
Contacts	0.000	High
Dossier age	0.002	Middle
Center-south	0.090	Low
Island	0.104	Low
North	0.618	Low

Table 14 Recovery Rate – Linear Model

<sup>9</sup> The model is consistent: at a more technical level: the F-statistic yields 13.12 to be compared with the critical f-value based on 6 and 619 degrees of freedom (over 120 DF the value is 1,774) to confirm that the model is statistically significant.

Macro region	%dossier	% Exposure	Real rec. rate	Fitted rec. rate
<b>North</b>	31.4%	33.1%	63.77%	65.47%
<b>Center North</b>	10.9%	10.6%	66.91%	64.60%
<b>Center South</b>	42.9%	41.1%	61.11%	59.67%
<b>Island</b>	14.8%	15.3%	58.60%	58.16%
<b>TOTAL</b>	21,234	22,900,489	62.6%	62.34%

Table 15 Comparison of Recovery Rate Outcome

The overall forecast for the Recovery Rate (RGR) is quite good (62.34% vs 62.6%). The match holds also at geographic area level, with the largest gap occurring for Centre-North where the model underestimates recovery rate by 2.3% on average.

## 10 Results

In order to measure the overall performance of the framework presented, recovery estimates at dossier level are summed up over the aggregating perimeter (macro-area) and then compared to the amount effectively recovered on the same perimeter,

$$(8) \quad \widehat{ER} = \sum_1^N E_i * I_i * \widehat{RgR}_i$$

$$(9) \quad I_i = \begin{cases} 1 & \text{if } PR \geq \text{cut-off} \\ 0 & \text{if } PR < \text{cut-off} \end{cases}$$

where the exposure of the given dossier selected by the recovery event ( $i$ ) at granular level - a non-linear function of the estimated probability of recovery – is multiplied by the estimated recovery rate for the given dossier ( $\widehat{RgR}_i$ ). This is different from the average recovered amount calculated through the average recovery rate over the given perimeter

$$(10) \quad \overline{ER} = \sum_{i \in \text{real rec}} E_i * \widehat{RgR}_i$$

as no exposure weighting is involved in the recovery rate estimate.

Hence the weighted average recovery rate at aggregated level, must be calculated from the expected recovered amount at granular level divided by their corresponding exposure.

$$(11) \quad \widehat{RgR} = \frac{\sum_{fit} E_i * I_i * \widehat{RgR}_i}{\sum_{fit} E_i * I_i}$$

It is worth mentioning that the average probability of recovery (RP) in a given perimeter is the ratio of dossiers recovered real or estimated

$$(12) \quad \overline{RP}_{real/fit} = 1/N \sum_{i=1}^N \widehat{I}_i$$

The results in terms of recovery parameters are summarized in the following table, confirming a good fit at aggregated level. Also the average probability of recovery is sufficiently precise, with the RF doing better than the Logit.

Macro region	Real		RgR%			PR%		
	Dossier #	Recovery/000	Real	Fitted	Real #	Real %	Logit %	RF optimal
<b>North</b>	6,665	166,6	63.8%	65.5%	211	3.2%	3.1%	3.1%
<b>Center North</b>	2,308	64,9	66.9%	64.6%	87	3.8%	4.6%	3.2%
<b>Center South</b>	9,109	154,2	61.1%	59.7%	246	2.7%	1.7%	2.5%
<b>Island</b>	3.152	50,8	58.6%	58.2%	82	2.6%	1.4%	2.8%
<b>TOTAL</b>	<b>21,234</b>	<b>436,5</b>	<b>62.6%</b>	<b>62.3%</b>	<b>626</b>	<b>2.9%</b>	<b>2.4%</b>	<b>2.8%</b>

Table 16 Fitted vs real recovery parameters summary

When a specific cut-off is defined for each approach, the resulting PR at macro-region level (2.8% RF optimal) gets very close to the real one (2.9%) and presents a more regular behaviour compared to that of the logistic approach which is also less precise (2.4%).

In terms of amounts recovered, the results are summarized in the following table, which compares the different components. Notice that the expression ‘exp\*Real PR’ next to exposure of effectively recovered (Real), represents an average based on the real probability of recovery defined at macro-region level. The two amounts differ as the last is not weighted for the relative exposure.

Macro region	amounts /000		Exp of recovered <sup>10</sup> /000				Tot recovery /000			
	Exposure	%Exp	Real	exp * Real RP	Logit	RF optimal	Real	%Real	Logit	RF optimal
North	7,572	33.1%	261.2	239.7	344,1	263,4	166,6	38.2%	229,9	160,6
Ctr North	2,433	10.6%	97.0	91.7	148,3	80,1	64,9	14.9%	97,2	46,7
Ctr South	9,402	41.0%	252.4	253.9	245,5	261,1	154,2	35.3%	140,6	135,9
Iseland	3,494	15.3%	86.7	90.9	93,9	144,3	50,8	11.6%	24,7	45,9
TOTAL	22,900	100%	697,3	676,2	831.7	748.8	436.5	100%	492.4	389,1

Table 17 Fitted vs real recovery amount summary

As for the recovery rate, the exposure of recovered dossiers estimated with the RF approach, driving the total recovery, is closer to the real one with respect to the Logit (748.8 vs 831.7 €). Considering the amounts recovered instead of the dossier exposure, the RF approach undershoots the real amount, but still performs slightly better than the Logit.

## Conclusions

This work analyzed the collection profiles in a portfolio of retail unsecured bills managed by a servicer, with the purpose of modelling the recovery capabilities on the basis of the limited information available on location and personal characteristics of the debtors. The exercise compares a RF approach to a classical Logistic estimate, using extended Confusion tables, besides other metrics, to compare their performances. To avoid averaging the recovery rate over a high number of non-recovery events, estimates are based on a framework that separates the recognition of recovery cases from the estimate of the recovery rate. The two components are analyzed separately. Although, in statistical terms, the performance of the model is generally low, mainly due to the very limited information available, at aggregate level the results are surprising. When the RF approach is used to identify recovery events, the results confirm the model to be robust with performance slightly better, than those of the standard Logistic, and with no need to define any specific conditions on explicative variables. Finally, the framework is suitable to estimate the portfolio value, both in absolute and relative terms, i.e. compared to a known portfolio. While there is a handful list of artificial intelligence methods which require large datasets, this work could be extended by applying some other machine learning technics, including support vector machine (SVM). As a classifier, SVM relies on the concept of distance between different points, a situation very similar to the characteristics of the data set considered in this work. However, as reported by (Wyner, et al., 2017), RF remains most likely the best classifier,.

Furthermore, the authors propose to combine Adaboost method with RF, using RF as the weak learner in the process for selecting the high weight instances during the boosting process. As reported by (Thongkam, et al., 2008), the proposed method outperforms a single classifier and other combined classifiers for the breast cancer survivability prediction.

Massimiliano Zanoni

## Bibliography

- Bellotti, A., Brigo, D., Gambetti, P. & Vrins, F., 2019. Forecasting recovery rates on non-performing loans with machine learning. s.l., s.n.
- Breiman, L., 1996. Out-of-Bag Estimation, s.l.: Statistics Department, University of California, Berkeley.
- Breiman, L., 2001. Random Forests, s.l.: Statistics Department, University of California, Berkeley.
- Ciavoliello, L. G. et al., 2016. What's the value of NPLs?. s.l., Banca D'Italia, Eurosistema.
- Fawcett, T., 2006. An introduction to ROC analysis. Pattern recognition letters.
- Gareth, J., Witten, D., Hastie, T. & Tibshirani, R., 2013. An Introduction to Statistical Learning with Applications in R. s.l.:Springer.
- Grömping, U., 2009. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. The American Statistician.
- Hastie, T., Tibshirani, R. & Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second edition. s.l.:Springer.
- Khieu, H., Mullineaux, D. J. & Yi, H.-C., 2012. The Determinants of Bank Loan Recovery Rates. Journal of Banking and Finance.
- Thongkam, J., Xu, G. & Zhang, Y., 2008. AdaBoost algorithm with random forests for predicting breast cancer survivability. s.l., International Joint Conference on Neural Networks (IJCNN).
- Wyner, A. J., Olson, M., Bleich, J. & Mease, D., 2017. Explaining the Success of AdaBoost and Random Forests as. Journal of Machine Learning Research, Volume 18.
- Youden, W., 1950. Index for rating diagnostic tests. Cancer.

<sup>10</sup> The exposure of recovered dossiers in any model is the sum over dossier recovered according to the specific model and cut-off.

# Stima prospettica delle misure finanziarie e di rischio mediante reti neurali dinamiche: un'applicazione al mercato statunitense

di Carlo Decherchi (Banca Carige) e Pier Giuseppe Giribone (Banca Carige, Università di Genova)

Articolo sottoposto a doppio referaggio anonimo, pervenuto in data 10/02/2020 e accettato il 24/03/2020

## Abstract

Lo scopo del presente articolo è quello di proporre l'impiego di una tecnologia basata su reti neurali dinamiche in grado di produrre stime prospettiche finanziarie finalizzate al calcolo di misure di rischio, quali ad esempio, l'Expected Shortfall (CVar). Lo studio, incentrato sul mercato statunitense, può essere suddiviso in quattro parti:

- la prima individua ed analizza i sedici indici rappresentativi del mercato statunitense dal 1999 al 2019. All'interno del periodo temporale di selezione vengono definiti tre sotto-periodi corrispondenti a differenti periodi storici considerati finanziariamente rilevanti: 1999-2006, 2006-2012 e 2012-2019.
- A partire dal set completo degli indici, nella seconda parte, si applica un'analisi di portafoglio al fine di determinare i pesi di ciascun asset nei periodi temporali di riferimento. Tale studio è propedeutico ad individuare, a partire dal campione complessivo di indici, quali tra questi risultano significativi in termini di asset allocation.
- Nella terza parte vengono presentati i principi di funzionamento delle reti neurali dinamiche NAR e NARX utilizzate come predittori per il sottoinsieme di indici risultati significativi nell'analisi quantitativa di portafoglio. Vengono pertanto fornite le misure di performance ottenute dalla rete durante la fase di addestramento (*in-sample*) e durante la fase di back-testing (*out-of-sample*).
- Una volta sicuri dell'affidabilità statistica ed econometrica dello strumento previsionale, nell'ultima parte del paper vengono calcolate le stime prospettiche dei valori degli indici, a partire dalle quali sono state stimate le misure finanziarie e di rischio in un'ottica forward-looking.

The purpose of this article is to explain how a technology based on dynamic neural networks is used for prices forecasting in order to calculate risk measures, such as the Expected Shortfall (CVar).

The paper is focused on US market and is divided into four parts:

- First section identifying and analyzing the sixteen most representative indexes of the US market from 1999 to 2019. Three sub-periods are thus defined, corresponding to financial relevant time-windows.
- In the second part, starting from the complete set of indexes, a portfolio analysis has been carried out in order to determine the weights of each asset in the reference time periods. This study is preliminary to identifying which of these are significant in terms of asset allocation.
- Third part discusses the working principles of NAR and NARX dynamic neural networks used as predictors for the subset of the more significant indexes. Moreover the performance measures obtained by the forecaster during the training phase (*in-sample*) and during the back-testing phase (*out-of-sample*) are provided.
- Once assessed of the statistical and econometric reliability of the forecasting instrument, the paper focuses on the future values of the indexes. Starting from these predictions, financial and risk measures have been estimated in a forward-looking perspective.

Disclaimer: Si fa presente che, essendo il presente documento redatto esclusivamente per scopi scientifici di ricerca, non vuole in alcun modo fornire al lettore informazioni di natura commerciale o finalizzate all'investimento. Il fine dichiarato del paper è quello di proporre una metodologia quantitativa robusta di analisi delle misure finanziarie e di rischio tramite metodologie di Machine Learning. L'approccio scientifico e quantitativo con il quale si sono condotte le simulazioni prospettiche dei livelli di prezzo degli indici considerati è proprio degli autori e non vincola in alcun modo le view delle Istituzioni di appartenenza.

## Key Words:

Artificial Neural Network (ANN), Dynamic Neural Network (DNN), Nonlinear Autoregressive (NAR) ANN, Nonlinear Autoregressive with exogenous variables (NARX) ANN, multi-step ahead prediction, CVar

## 1 Selezione ed analisi degli indici rappresentativi del mercato statunitense

Sono stati selezionati 16 indici settoriali del mercato statunitense, che risultano preferibili ad ETF o fondi comuni ai fini di una analisi statistica, in quanto non influenzati da alcun tipo di gestione (passiva o attiva) ed i ritorni non subiscono alcun tipo di distorsione conseguente ai costi di gestione, di entrata e di uscita.

Il mercato USA è stato preferito a quello Europeo per evitare problematiche riconducibili a distorsioni create dall'eterogeneità del tessuto economico dell'UE, oltre che ai differenti livelli di tassazione che senz'altro inficiano anche il livello dei prezzi espresso dal mercato.

Discorso analogo è valido per la definizione di un teorico tasso privo di rischio comune ai diversi paesi, che risulterebbe soggetto a semplificazioni e distorsioni.

Gli indici sono stati scelti con l'intento di creare una visione più completa possibile dell'economia nel suo complesso ed il loro valore viene calcolato con una metodologia analoga a quella utilizzata per lo Standard & Poor's 500, che consiste nella ponderazione dei ritorni per il livello di capitalizzazione delle società presenti dell'indice.

Impiegando l'info-provider Bloomberg®, sono stati creati 16 vettori dei prezzi degli indici, rilevati con cadenza giornaliera, dal 20/10/1999 al 15/10/2019, di cui viene anche riportato il codice identificativo Bloomberg®:

automotive (S5AUOC Index);  
bancario (S5BANKX Index);  
chimico e farmaceutico (S5CHEM Index);  
costruzioni e del real estate (S5CSTMX Index);  
mercato dell'energia (SPTRENRS Index);  
intrattenimento e mass media (S5MEDA Index);  
settore finanziario (S5FINL Index);  
food & beverage (S5FDBT Index);  
healthcare (S5HLTH Index);  
produzione industriale (S5INDU Index);  
mercato assicurativo (S5INSU Index);  
oil & gas (S5OILG Index);  
beni di consumo retail (S5RETL Index);  
tecnologico (S5INFT Index);  
telecomunicazioni (S5TELS Index);  
utilities (S5UTIL Index).

Dopo un trattamento preliminare dei dati grezzi, è stato costruito il “dataset” definitivo, suddiviso in intervalli temporali dal 20/10/1999 al 09/10/2019, per un totale di 5025 osservazioni per singolo vettore [9].

A partire da tale campione sono stati calcolati i rendimenti giornalieri (5024 osservazioni) e mensili (240 osservazioni), che saranno oggetto di analisi per la definizione delle frontiere efficienti e per il calcolo del CVar prospettico.

All'interno del periodo temporale di selezione vengono definiti tre sotto-periodi, corrispondenti a differenti periodi storici considerati rilevanti:

- Primo periodo: dal 20/10/1999 al 01/09/2006, ivi compresa la bolla “dotcom” ed il periodo pre-recessione;
- Secondo periodo: dal 01/09/2006 al 01/01/2012, situazione recessiva globale, iniziata negli USA nel 2007 e nei fatti perdurata fino a tutto il 2011, fondamentale per lo studio delle allocazioni efficienti in periodi di crisi;
- Terzo periodo: dal 01/01/2012 al 09/10/2019, fase di crescita economica e, da circa metà del 2018, valutazione di un secondo possibile periodo pre-crisi.

La loro definizione deriva da una rigorosa verifica della struttura a termine dei tassi di interesse dei titoli di stato statunitensi, in cui si è cercato di individuare i periodi caratterizzati da una prolungata ed evidente inversione della curva stessa, segnale assunto in letteratura come previsore di probabili periodi recessivi ed indicatore del “sentiment” sui mercati, quindi del livello di fiducia degli investitori circa l'andamento futuro dell'economia.

Per lo studio vengono presi in esame i tassi di interesse a 3 mesi, a 2 anni, a 10 anni e a 20 anni, in modo da avere una comprensione completa della curva [6].

Il test condotto consiste nella creazione di una variabile di tipo “dummy”, in cui la ratio applicata prevede che il tasso di interesse a 3 mesi, quindi di breve periodo, sia simultaneamente maggiore degli altri tre. In caso positivo, il test restituirà valore 1, in caso contrario 0. Il risultato così raggiunto avvalora la scelta dei 3 sotto-periodi.

La caratterizzazione finanziaria, in termini di rendimenti mensili medi e correlazione per i quattro periodi temporali (1999-2006, 2006-2012, 2012-2019 e 1999-2019), è riportata, rispettivamente, nei Panel 1 e 2.

## 2 Analisi del portafoglio

A partire dal set completo dei 16 indici selezionati nel paragrafo 1, si applica un'analisi di portafoglio al fine di determinare i pesi assunti da ciascun asset nei quattro periodi temporali di riferimento [9], [12].

### 2.1 Primo periodo temporale: 1999-2006

Partendo dall'osservazione del primo periodo, complice anche la bolla speculativa nel mercato tecnologico, la frontiera generata presenta un portafoglio di tangenza, quindi con *Sharpe Ratio* massimizzato[16], composto quasi esclusivamente da titoli azionari del settore energetico (Panel 3).

In prima battuta, si nota che il tasso privo di rischio, pari ad un valore medio annuo del 3,92%, differisce solamente per l'1% rispetto al rendimento atteso del portafoglio a minima varianza, come riassunto dai valori nella Tabella 1.

ANALYTICS 1999 - 2006	MV	Tangenza	EW
E[r]	4,53%	16,79%	2,62%
$\sigma$	10,99%	11,39%	13,75%
SR	0,0555	1,1299	-0,0945
MR	0,4122	1,4741	0,1905

Tabella 1. Analytics 1999-2006

Insieme agli indicatori classici, viene definito il “Market Ratio” (MR), inteso come il rapporto tra E[r] e  $\sigma$ , che non tiene conto dell’asset privo di rischio ed esprime l’efficienza “pura” del portafoglio, restituendo valori per definizione maggiori rispetto allo SR, il quale invece è proprio di uno scenario realistico di mercato.

In termini di allocazione, vengono riportati gli istogrammi dei pesi, rispettivamente per i portafogli a minima varianza e di tangenza nei Panel 4 e 5.

Nel Panel 5 si evidenzia come la strategia più efficiente consista nell’investire la quasi totalità delle risorse nel settore energetico, con un peso del 74,79%, a fronte di un residuale 25,21% nel settore dell’edilizia, reso redditizio dal boom immobiliare pre-recessivo.

Nel Panel 4 viene invece rappresentata la strategia in grado di minimizzare la varianza totale di portafoglio, la cui realizzazione richiede invece un’allocazione maggiormente articolata, passando da soli 2 asset a 9.

Per valutare complessivamente i risultati della strategia lungo il periodo di osservazione, viene mostrato l’andamento grafico dei due portafogli ottimali, affiancati dal benchmark di mercato S&P500 Total Return. Inoltre, si considera l’ipotesi di reinvestimento al tasso privo di rischio, con valore di partenza pari a 100 (Panel 6).

Da sottolineare, oltre all’evidente sovra-performance del portafoglio di tangenza, le differenze in termini di volatilità.

Come osservato nel Panel 6, la strategia priva di rischio, oltre a garantire rendimenti migliori rispetto al benchmark di mercato da fine anno 2000, restituisce performance migliori, fino alla fine del 2003, rispetto alla strategia di tangenza e fino a metà del 2005 per l’approccio di minima varianza. Successivamente a tale periodo temporale, entrambe le strategie forniscono invece risultati migliori dell’investimento in titoli di stato.

## 2.1 Secondo periodo temporale: 2006-2012

Nel secondo periodo, in fase recessiva dell’economia globale, si è reso necessario introdurre come diciassettesimo asset l’oro per definire, come mostrato nel Panel 3, una frontiera efficiente capace di creare migliori livelli di diversificazione. In questo sotto-periodo, il tasso privo di rischio assume un valore pari al 2,8% medio annuo.

Nella Tabella 2 è possibile notare l’impatto positivo della commodity all’interno del portafoglio, portando la strategia di tangenza a presentare uno SR pari a 0,7659, quale segnale di alti livelli di efficienza, nonostante il periodo di crisi economica globale. Rispetto al periodo precedente, le strategie risentono dell’incertezza di mercato, che si traduce in una maggiore volatilità.

ANALYTICS 2006 - 2012	MV	Tangenza	EW
E[r]	7,55%	15,1%	-1,47%
$\sigma$	11,55%	16,06%	20,41%
SR	0,4113	0,7659	-0,2297
MR	0,6537	0,9402	-0,0702

Tabella 2. Analytics 2006-2012

Dalle allocazioni mostrate nei Panel 4 e 5, emerge un elemento degno di nota: l’oro contribuisce positivamente e in parte preponderante al rendimento della strategia di tangenza, come lecito aspettarsi, data la natura stessa dell’asset. Si noti che tale comportamento si affianca il settore del cibo e delle bevande (F&B), con un peso del 28,91%.

Se si guarda alla strategia a varianza minimizzata, quest’ultimo settore di mercato restituisce un peso pari al 43,03%, mentre il peso dell’oro si ferma al 23,59%.

Questo aspetto risulta particolarmente importante in quanto il settore di cibi e bevande ritrova nuovo vigore durante il periodo recessivo, dettato dal reddito medio pro-capite in calo, che sposta quindi i consumi verso beni più accessibili.

Inoltre, questo fenomeno è stato accelerato dal rapido cambiamento dello stile di vita del cittadino americano medio, che predilige prodotti a buon mercato caratterizzati da una fruibilità immediata.

Dall'osservazione del Panel 6, si evince che il portafoglio a minima varianza garantisce anch'esso un rendimento totale del 48,6%, pur trovandosi in una fase economica recessiva. La strategia di tangenza riconferma i risultati migliori, con una performance, in termini di rendimento atteso, pari al 120,95%, valore che, come affermato, è giustificato dal bene anticiclico presente nel paniere.

### 2.3 Terzo periodo temporale: 2012-2019

Nel terzo periodo, quindi in fase di crescita economica, si assiste a risultati decisamente diversi da quelli mostrati fino ad ora. A seguito alle politiche monetarie espansive adattate dalla FED e del “Quantitative Easing” in Europa, in aggiunta alle novità normative stringenti per il sistema finanziario nel suo complesso, alla crescita esponenziale del settore tecnologico e a una rinnovata attenzione al tema del cambiamento climatico, il tessuto economico ha subito profondi mutamenti che perdurano tutt'oggi. In questo periodo, le politiche monetarie espansive appena citate hanno creato forti ribassi nei tassi di interesse, nel tentativo di stimolare gli investimenti e la circolazione della moneta. Ne è conseguito un tasso privo di rischio minore in assoluto, con un rendimento medio annuo di 1,45%.

La frontiera efficiente del periodo rappresentata nel Panel 3 mostra graficamente le conseguenze di un tasso privo di rischio ai minimi storici. L'intercetta della CAL fa sì che questa abbia una elevata pendenza, col chiaro intento, da parte delle banche centrali, di disincentivare l'immobilizzazione dei capitali favorendo gli investimenti in attività reali, con impatti diretti sulla crescita economica e sui rendimenti dei mercati. La Tabella 3 riassume le performance.

ANALYTICS 2012 - 2019	MV	Tangenza	EW
E[r]	9,29%	14,99%	9,65%
$\sigma$	8,81%	10,92%	11,54%
SR	0,8899	1,2399	0,7106
MR	1,0544	1,3727	0,8362

Tabella 3. Analytics 2012-2019

Le politiche monetarie espansive delle banche centrali hanno adottato politiche di ribasso dei tassi d'interesse, incentivando gli operatori ad investire nel mercato azionario e obbligazionario societario (“corporate”), per sostenere sia la circolazione dei capitali sia per dare nuova linfa al sistema economico in lieve ripresa post-crisi.

In aggiunta, ne consegue un deciso calo della disoccupazione e ad un contestuale aumento dei consumi, del PIL e dell'inflazione.

Nei Panel 4 e 5 viene mostrato che in entrambe le strategie spicca per efficienza il settore delle “Utilities”, con un risultato particolarmente accentuato nella strategia di minima varianza. Trattandosi di aziende che forniscono servizi essenziali, come acqua, gas, elettricità e perciò produttrici di beni a domanda tendenzialmente anelastica rispetto al prezzo. Per questo, indipendentemente dalla fase del ciclo economico e dalla velocità con cui l'economia cresce, tale settore risulta tra i più stabili del mercato azionario. Infatti, il suo rendimento medio risulta inferiore rispetto a quello di altri settori, come per esempio quello tecnologico, a fronte di una minore volatilità storica. Per cui, complici nuovi metodi più efficienti di produzione, soprattutto dell'energia elettrica, la nascita della “green economy” e la sua repentina crescita nel mercato e la bassa volatilità (che per natura caratterizza i titoli delle società che compongono tale indice settoriale), si può osservare un peso rilevante nel portafoglio di tangenza e predominante in quello a minima varianza, nei quali pesa rispettivamente per il 28,49% e per il 39,79%. In termini di risultati finali, nel Panel 6 viene mostrato come il portafoglio di tangenza riesca nuovamente a battere il benchmark di mercato, seppur in misura minore, data la crescita complessiva del mercato sintetizzata dall'indice stesso. La strategia a minima varianza raggiunge il suo scopo, mitigando la volatilità ed ottenendo al contempo una crescita di periodo del 107,07%. Di fatti, rispetto alla strategia di tangenza, il calo di volatilità risulta pari a 211 punti base, differenziale in termini di rischio, in linea con quanto osservato nei risultati precedenti.

### 2.4 Serie storica complessiva: 1999-2019

Nel periodo complessivo vengono infine analizzati 20 anni di osservazioni, quindi comprendendo un ciclo economico completo. Il tasso d'interesse annuale medio si assesta al 2,72%. Come osservabile nel Panel 3, i settori che hanno registrato i valori peggiori sono quelli delle telecomunicazioni e delle automobili.

Dalla Tabella 4 emerge come aspetto principale che, nel lungo periodo, lo SR diminuisce notevolmente, in quanto calcolato utilizzando i ritorni medi registrati in tutte le fasi del ciclo economico, dalla fase espansiva alla recessiva. I ritorni medi risultano necessariamente più bassi rispetto ai ritorni delle fasi espansive dell'economia, accompagnati da una maggiore volatilità.

ANALYTICS 1999 - 2019	MV	Tangenza	EW
E[r]	5,5%	8,51%	3,97%
$\sigma$	11,2%	13,95%	15,45%
SR	0,2482	0,4151	0,0809
MR	0,4911	0,6101	0,2569

Tabella 4. Analytics 1999-2019

Il Panel 4 riconferma, anche in questo caso, come settore predominante il F&B, con un peso pari al 35,62%, seguito dal mercato della salute con il 22,86% e dal settore delle “utilities” col 22,7%. Il panel 5 mostra il duopolio, in termini di allocazione efficiente, del settore F&B e dei consumi retail.

Fra i settori analizzati, quelli alimentare, assieme ai consumi retail, settore HC ed “utilities”, risultano i più efficienti in termini allocativi in una strategia di lungo periodo. Infine, come sintetizzato nel Panel 6, si mostra come le strategie abbiano nuovamente ottenuto performance storiche migliori rispetto all’indice di mercato.

In base alle risultanze acquisite è possibile identificare quei settori che, nei diversi periodi, hanno contribuito alla creazione delle diverse strategie [16], [12]. In particolare, visti i risultati, si decide di escludere dalle successive analisi i settori automotive, dell’industria chimica, dell’intrattenimento, della produzione industriale e finanziario.

Selezionati a partire dal campione complessivo gli indici risultati significativi in termini di asset allocation, si procede, per tale sottosinsieme, a calcolare le proiezioni impiegando come strumento di forecasting le reti neurali dinamiche di tipo NAR (Non-Linear Auto Regressive) e NARX (Non-Linear Auto Regressive with Exogenous Variables) [10].

### 3 Le reti neurali dinamiche NAR e NARX come strumento previsionale e relative misure di performance

In questa sezione vengono presentati sinteticamente i principi di funzionamento delle reti neurali dinamiche e le relative procedure di addestramento [15], [1]. Successivamente, vengono presentate le misure di performance impiegate per comprovare la qualità della previsione in applicazione al back-testing finalizzato alla validazione della rete [14]. Per approfondimenti teorici più specifici inerenti le reti neurali statiche e dinamiche, si rimanda alla letteratura scientifica di settore citata in bibliografia [4].

#### 3.1 Funzionamento di una rete neurale dinamica di tipo NAR e NARX

NAR e NARX fanno parte della classe di reti neurali dinamiche auto-regressive non lineari (“*Non Linear Autoregressive network*”), impiegate anche a fini predittivi, con la capacità di prevedere i valori futuri di una serie storica a partire dai suoi valori passati [8]. Le reti neurali dinamiche presentano approcci computazionali basati sul “*feedback delay*” (o ritardo), così che gli output non dipendano solamente dagli input, ma anche dagli output e dagli stati ottenuti in precedenza [3].

Si indica con  $y(t)$  una serie storica di  $n$  osservazioni:  $y(t) = f(y_{t-1}, y_{t-2}, \dots, y_{t-n})$ , in cui i regressori per il valore presente vengono assegnati dai valori passati  $y_{t-1}, y_{t-2}, \dots, y_{t-n}$  [17]. Un tipo di analisi di  $y(t)$  consiste nell’applicazione di reti neurali di tipo ricorrente. Questo approccio genera un processo con diversi gradi di memoria, rendendolo particolarmente idoneo agli obiettivi di forecasting, dal momento che le reti possono essere addestrate a distinguere trend temporali. Graficamente, l’architettura di un singolo strato risulterà del tipo rappresentato nella Figura 1 [2]. In particolare:  $p(t)$  è il vettore input della rete, che viene ponderato per  $IW$  (“*input weight*”);  $n(t)$  invece rappresenta l’uscita dell’ “*hidden layer*”, definito come la somma pesata del vettore d’ingresso, moltiplicato per il rispettivo peso, unitamente al contributo del segnale fornito dai regressori della serie temporale, moltiplicati per il rispettivo peso  $LW$  (“*layer weight*”);  $a(t)$ , che risulta essere l’uscita dell’output layer della rete, caratterizzato quindi da un blocco di trasferimento lineare, risulta funzione dell’input corrente  $p(t)$  e del precedente output  $a(t-1)$ , regolato dal blocco  $D$  (*delay*). Quest’ultimo è, a sua volta, espresso in funzione di  $p(t-1)$  e  $a(t-2)$  e così via.

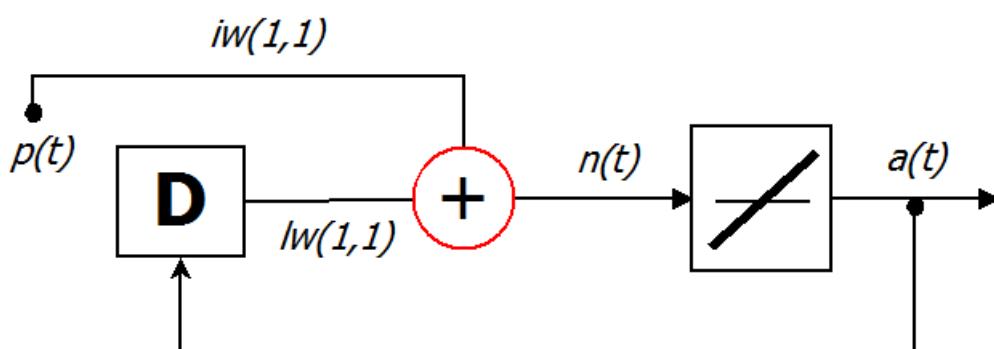


Figura 1. Schematizzazione di una rete neurale ricorrente ad uno strato

Con la generalizzazione di questo tipo di rete si perviene ad una NARX, che segue le logiche descritte ed è caratterizzata da un'architettura più articolata, con conseguenti tempistiche di calcolo maggiori, per la presenza di variabili esogene.

La serie storica è anche in funzione di una variabile esogena:  $y(t) = f(y_{t-1}, y_{t-2}, \dots, y_{t-n_y}, u_{t-1}, u_{t-2}, \dots, u_{t-n_u})$ , dove viene introdotto il numero di ritardi da applicare sia alla variabile endogena (il livello dei prezzi),  $n_y$ , sia alla variabile esogena (ad esempio i volumi),  $n_u$  [10].

Come viene mostrato nella Figura 2, questa architettura di rete permette una flessibilità maggiore in relazione alla potenziale dinamica futura seguita dalla time-series.

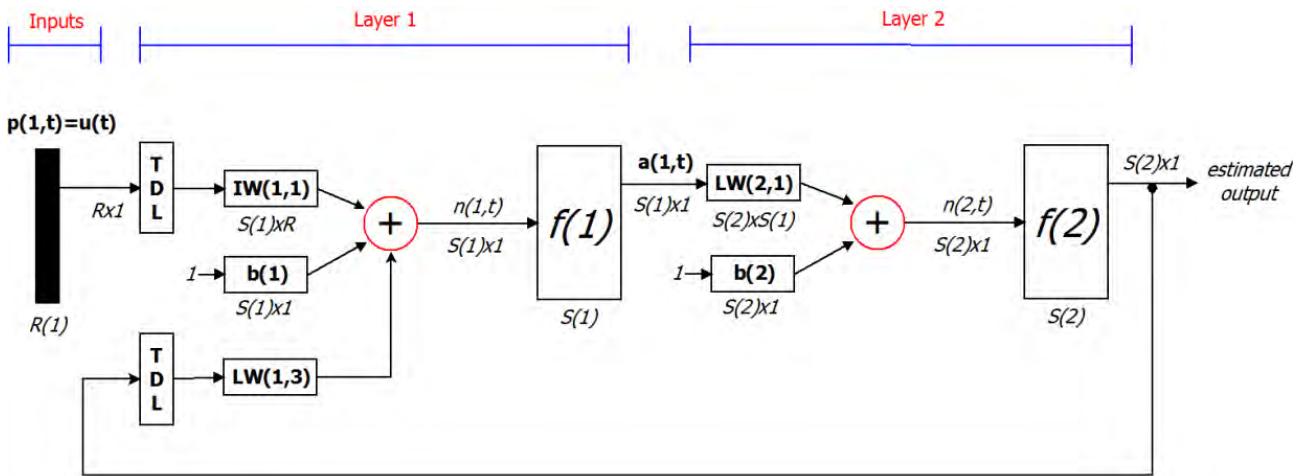


Figura 2. Esempio di architettura NARX a due strati nascosti

È necessario discriminare tra i pesi degli input ( $IW$ ) e i pesi degli output degli strati ( $LW$ ), in quanto ogni strato, tranne naturalmente l'ultimo, rappresenta il vettore input per quello immediatamente successivo, che trasmette il segnale generato tramite la propria funzione di trasferimento ( $f$ ). Analogamente alla configurazione precedente, il blocco che gestisce i ritardi è definito come  $TDL$  ("tapped delay")

A questi pesi, al fine di rendere maggiormente flessibile l'architettura della rete, è prevista la possibilità di aggiungere un ulteriore vettore di parametri chiamati bias e rappresentati in Figura 2 come  $b$ .

In analogia alle reti statiche, anche nelle reti dinamiche l'addestramento viene eseguito mediante la minimizzazione di una funzione errore, che quantifica la discrepanza tra il valore osservato e quello fornito dal modello, tipicamente espressa dalla MSE [7], [11].

La taratura dei parametri della rete ( $IW$ ,  $LW$  e  $b$ ) avviene nel corso di tale processo di ottimizzazione numerica [13]. Una rete NAR e NARX può essere considerata un predittore ragionevole, qualora sia dotato di un buon fitting statistico (misurato tipicamente con  $R^2$ ) e in assenza di auto-correlazione dell'errore.

### 3.2 Training delle reti

Per ogni singolo indice significativo selezionato nel paragrafo 2, la rete neurale dinamica è stata addestrata e calibrata in modo tale da:

- registrare valori prossimi a 1 di  $R^2$ , misura statistica che definisce l'andamento fedele della serie storica e a generalizzarne il comportamento;
- non osservare autocorrelazione degli errori, accertandosi che ogni "bin" calcolato non superi le soglie di tolleranza.

L'obiettivo primario risiede nella minimizzazione della già citata *loss function* (MSE) ed è riconosciuto come indicatore delle performance del processo di addestramento.

L'algoritmo base utilizzato nell'addestramento delle reti neurali è "gradient descent", caratterizzato dalla seguente formula di aggiornamento dei pesi:

$$w(i+1) = w(i) - \mu(i)g(i)$$

dove  $w(i+1)$  è il vettore dei pesi calcolati all'iterazione  $i+1$ ,  $w(i)$  il vettore dei pesi calcolati all'iterazione iniziale  $i$ ,  $\mu(i)$  è il tasso di apprendimento all'iterazione  $i$  e  $g(i)$  è il gradiente calcolato all'iterazione  $i$ .

Il processo utilizzato per il calcolo del gradiente  $g(i)$  è detto di "back-propagation", che consiste in un'analisi a ritroso, con lo spostamento degli errori dal vettore di uscita a quello di ingresso.

L'algoritmo implementato, spesso utilizzato nelle reti neurali dinamiche, è il Levenberg-Marquardt [13], che prevede la seguente formula di aggiornamento dei pesi:

$$w(i+1) = w(i) - [J^T J + \mu(i)I]^{-1} J^T$$

dove  $J$  è la matrice Jacobiana delle derivate prime degli errori rispetto ai pesi e ai bias,  $\mu(i)$  definisce l'ampiezza dello step dell'iterazione, che diminuirà man mano che ci si avvicina al minimo della funzione di performance.

Per ottenere quanto esposto, si è dovuto modulare sia le dimensioni dello strato nascosto, sia l'intervallo di ritardo di "feedback" [9].

Nella Tabella 5 vengono elencati, per le varie tipologie di indici, le diverse calibrature, unitamente ai livelli di  $R^2$  e MSE, ottenuti per la serie identificata dalla rete e per quella empirica della rete dinamica di tipo NAR.

I risultati mostrano livelli di  $R^2$  ottimali, tendenti a 1, che permettono alla rete di riprodurre fedelmente l'andamento della serie storica e di generalizzarne il comportamento nel futuro.

Il livello medio di MSE è pari a 98,9363, con un picco negativo registrato durante l'addestramento sull'indice di mercato di 770,4749, a fronte di quello migliore pari a 3,4795 sul settore delle telecomunicazioni.

La dimensione media dello strato nascosto, quindi del livello di complessità richiesto per raggiungere livelli accettabili di qualità, risulta pari a 4.

Infine, l'intervallo di "feedback" è stato scelto, come anticipato, con un valore simile a quello degli intervalli da simulare. Per questo motivo, i valori ottimali si assestano fra 150 e 200.

RISULTATI ADDESTRAMENTO	Dim. Strato nascosto	Intervallo di feedback	$R^2$	MSE
Banche	4	1:150	0,994	15,5946
Real Estate	6	1:150	0,9939	7,0696
Energia	6	1:150	0,9941	99,6395
F&B	4	1:150	0,9948	11,2750
HC	4	1:150	0,9947	29,0391
Assicurativo	4	1:150	0,994	14,7415
Oil&Gas	4	1:200	0,9939	41,2165
Cons. Retail	3	1:150	0,9948	128,0557
Tecnologia	3	1:200	0,9946	62,5506
Telecomun.	4	1:200	0,9929	3,4795
Utilities	4	1:200	0,9944	4,0987
S&P500 TR	2	1:200	0,9948	770,4749

Tabella 5. Risultati dell'addestramento della rete NAR nei diversi settori

Nella Tabella 6 vengono mostrate le configurazioni utilizzate per l'apprendimento della rete NARX.

RISULTATI ADDESTRAMENTO	Dim. Strato nascosto	Intervallo di feedback	$R^2$	MSE
Banche	4	1:150	0,9942	15,3249
Real Estate	2	1:170	0,994	7,2876
Energia	6	1:150	0,9943	93,9782
F&B	3	1:150	0,9948	11,8384
HC	4	1:150	0,9948	30,2630
Assicurativo	2	1:200	0,9939	13,9342
Oil&Gas	2	1:130	0,994	42,9955
Cons. Retail	4	1:150	0,9948	125,0059
Tecnologia	3	1:160	0,9948	70,1123
Telecomun.	4	1:130	0,9932	3,7823
Utilities	5	1:150	0,9942	4,1179
S&P500 TR	2	1:140	0,9948	761,1433

Tabella 6. Risultati dell'addestramento della rete NARX nei diversi settori

Anche in questo frangente, la rete restituisce livelli di  $R^2$  tendenti a 1, confermando la capacità di adattamento della rete ai dati.

Il livello medio di MSE è pari a 98,3153, di poco inferiore rispetto a quello della rete NAR, con un picco negativo, anche in questo caso, sull'indice S&P500 di 761,1433, giustificabile dalla magnitudo dei livelli dell'indice, a fronte del migliore pari a 3,7823, nuovamente registrato sul settore delle telecomunicazioni.

La dimensione media dello strato nascosto risulta pari a 3,4, quindi inferiore rispetto alla tipologia precedente, perseguito il raggiungimento di risultati migliori. L'intervallo di "feedback" presenta valori ottimi fra 130 ed i 200.

Nonostante la riduzione dell'ampiezza dell'intervallo e della dimensione dello strato nascosto, si osserva una maggiore efficienza della rete NARX.

In aggiunta, ciò permette di ridurre notevolmente i tempi di calcolo dell'addestramento della stessa e, di conseguenza, di simulazione dei prezzi futuri.

### 3.3 Misure di performance e back-testing

Il test out-of-sample viene condotto sulle ultime 125 osservazioni empiriche e sulle 125 osservazioni simulate dalla rete a partire dall'osservazione numero 4900. L'arco temporale in esame va dunque dal 11/04/2019 al 09/10/2019.

Per farlo vengono analizzati tre differenti misure del grado di dispersione [5]. In primis, RMSE ("Root mean square error", radice dell'errore quadratico medio):

$$RMSE = \sqrt{\frac{\sum_i^n \varepsilon_i^2}{n}}$$

come secondo indicatore si utilizza il MAE ("Mean absolute error", errore assoluto medio):

$$MAE = \frac{\sum_i^n |\varepsilon_i|}{n}$$

infine, si utilizza lo MSEP ("Mean square error Percentual", errore quadratico medio percentuale):

$$MSEP = \frac{\sum_i^n \left(\frac{\varepsilon_i}{\hat{y}_i}\right)^2}{n}$$

in cui  $y_i$  rappresenta il valore per l'i-esima osservazione,  $\varepsilon_i$  è invece la differenza tra dato empirico  $y_i$  e teorico  $\hat{y}_i$  ed n la numerosità dei dati.

Nella Tabella 7 vengono riportati i risultati ottenuti applicando tale metodologia agli undici indici settoriali.

Partendo dalla valutazione dello MSEP, è possibile constatare che la rete sia stata in grado di cogliere, in modo statisticamente rilevante, l'andamento della serie storica di 9 indici su 11 considerati.

Ciò permette di generare un vettore previsionale verosimile, che tenga conto dell'andamento storico, se pur soggetto a qualche limitazione, superata dal modello NARX, che implementa anche la variabile esogena dei volumi.

Medesimo ragionamento è valido osservando i valori di RMSE, con valore medio pari a 42,4232, unitamente a MAE, con media pari a 35,1631.

La confrontabilità dell'errore tra diverse serie storiche, resa possibile principalmente dallo MSEP, indica un valore medio di quest'ultima di 0,0038, con eterogeneità dei valori, calcolata con la deviazione standard, pari a 0,0044, dal minimo ottenuto nell'addestramento sul benchmark pari a 3,9785e-04 al massimo del settore assicurativo, pari al 0,0129.

Il potere previsionale della rete NAR risulta quindi limitato soprattutto per previsioni nel medio-lungo termine, ma ciononostante rimane una tecnica ragionevolmente predittiva dei valori futuri, nel breve periodo, di una serie storica finanziaria.

Analoga procedura di validazione è stata condotta usando una rete NARX, impiegando come variabile esogena i volumi. Nella Tabella 8 è riportata l'analisi statistica.

ANALISI STATISTICA EMPIRICO VS PREVISTO NAR	RMSE	MAE	MSEP
Banche	16,1275	13,7919	0,0024
Real Estate	29,4260	25,6441	0,0125
Energia	49,3075	41,8394	0,0037
F&B	27,1965	23,9240	0,0016
HC	28,0989	22,228	7,3291e-04
Assicurativo	49,2315	45,92	0,0129
Oil&Gas	34,2001	26,3449	0,0048
Cons. Retail	92,241	73,3608	0,0015
Tecnologia	41,9159	35,4687	9,0170e-04
Telecomun.	3,5995	2,7423	4,9970e-04
Utilities	18,8849	14,0218	0,0035
S&P500 TR	118,8489	96,6709	3,9785e-04

Tabella 7. Performance out-of-sample delle rete NAR nei diversi settori

ANALISI STATISTICA EMPIRICO VS PREVISTO NARX	RMSE	MAE	MSEP
Banche	10,1048	8,3138	0,001
Real Estate	26,2790	20,5854	0,0094
Energia	32,8759	26,6274	0,0016
F&B	17,2736	14,7349	6,3699e-04
HC	28,1787	22,2903	7,5064e-04
Assicurativo	24,4833	22,376	0,0032
Oil&Gas	34,8313	27,5059	0,0053
Cons. Retail	72,3322	55,8991	9,2130e-04
Tecnologia	36,1599	27,6207	7,5267e-04
Telecomun.	31,5824	3,2010	6,9986e-04
Utilities	16,2219	13,4064	0,0026
S&P500 TR	220,8449	185,2784	0,0014

Tabella 8. Performance out-of-sample delle rete NARX nei diversi settori

Dai valori riassunti è desumibile un significativo miglioramento delle performance della rete in termini di apprendimento. Osservando, in prima battuta, il valore medio dello MSE, questo scende ad un valore pari a 0,0024, con contestuale calo della variabilità dei valori dell'indicatore a 0,0026. In questo frangente, la rete è stata in grado di cogliere, in modo statisticamente rilevante, l'andamento della serie storica.

#### 4 Stima prospettica delle misure finanziarie e di rischio mediante NARX

Essendosi rilevata la rete NARX più performante in sede di back-testing e validazione [9], si impiegherà per la stima prospettica dei livelli di prezzo nel periodo temporale, compreso tra l'ottobre 2019 e l'ottobre 2020. I valori prospettici si analizzeranno sia dal punto di vista finanziario, impiegando le tecniche tradizionali della teoria di portafoglio (come effettuato per i rendimenti storici nel paragrafo 1), sia dal punto di vista dell'analisi di rischio. Vengono quindi mostrati i risultati in termini di rischio e rendimento del vettore previsionale ottenuto dalla rete per i diversi settori, di cui vengono mostrati in Figura 3, come primi elementi di analisi, i rendimenti logaritmici medi mensili. La rete sembra dunque rinforzare la previsione di generale rallentamento dell'economia statunitense. Anche in questo scenario, l'impatto maggiore viene osservato nel settore degli immobili, seguito da quello delle "utilities". Gli unici settori con rendimenti positivi degni di nota risultano essere nuovamente il settore bancario, di oil&gas e dell'energia, essendo per propria natura, questi ultimi due, fortemente correlati.

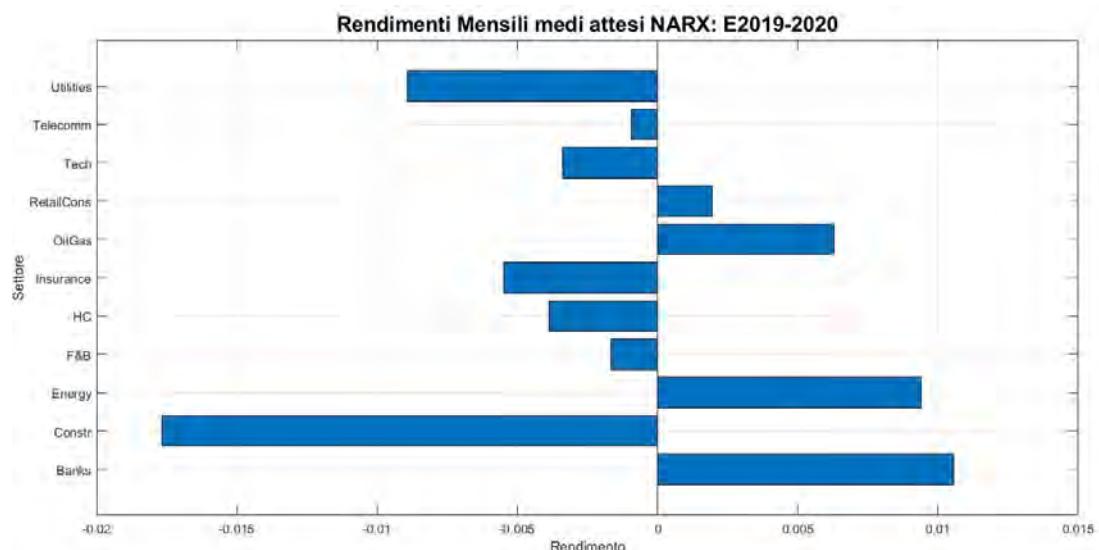


Figura 3. Rendimenti mensili medi previsti con rete NARX da Ottobre 2019 ad Ottobre 2020

In Figura 4 viene mostrata la "heatmap" delle correlazioni prevista dalla rete NARX.

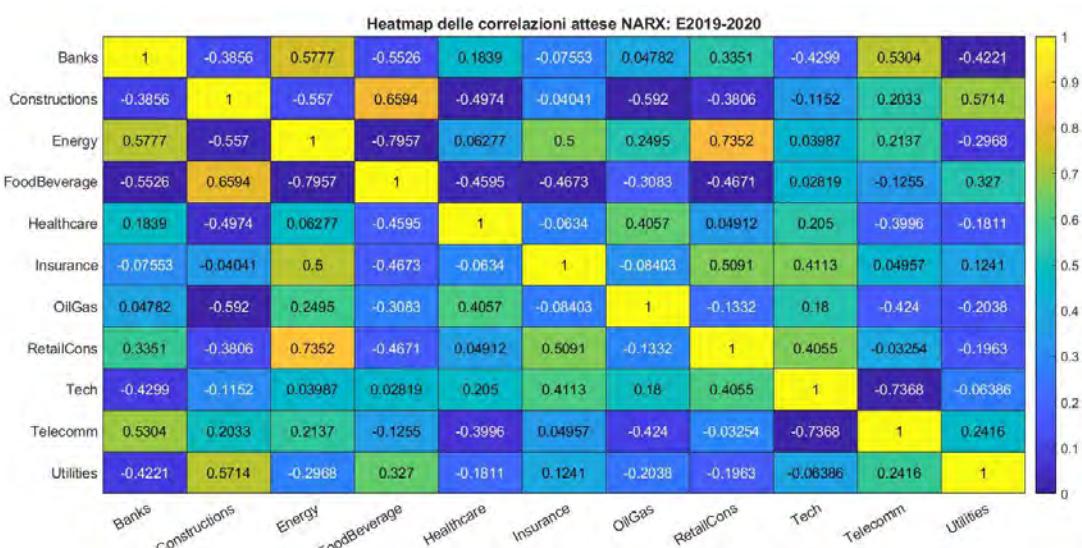


Figura 4. Heatmap delle correlazioni attese con rete NARX da Ottobre 2019 ad Ottobre 2020

I maggiori livelli di volatilità ed i risultati eterogenei di breve periodo, determinano livelli di correlazione in media minori dei precedenti. La struttura della frontiera efficiente di portafoglio attesa, presenta l'architettura mostrata in Figura 5.

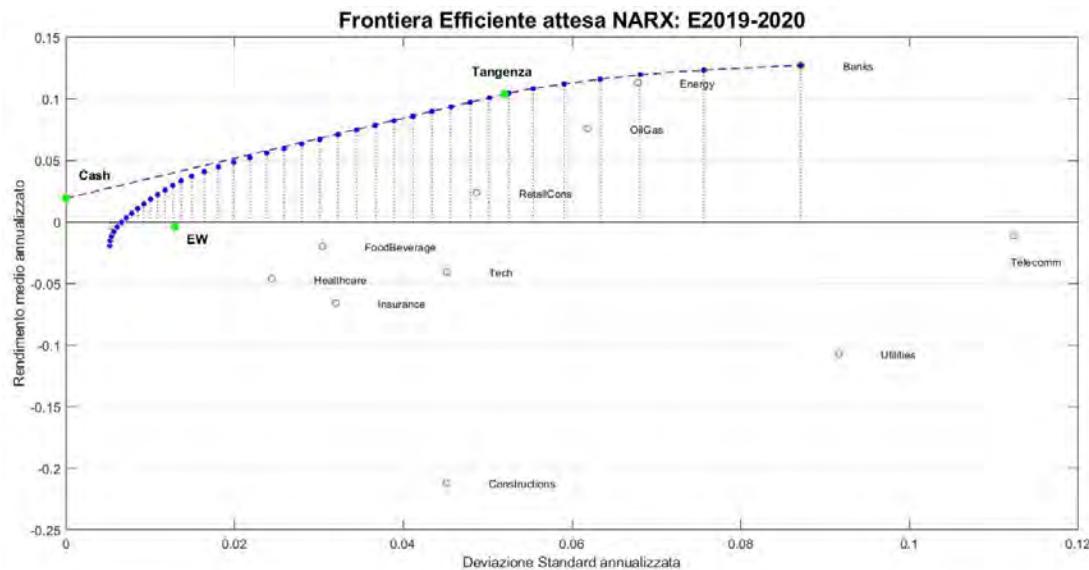


Figura 5. Frontiera efficiente attesa con rete NARX da Ottobre 2019 ad Ottobre 2020

Lo scenario individuato esprime le caratteristiche tipiche di un periodo di incertezza nei mercati finanziari. In particolare, la strategia di minima varianza registra un rendimento inferiore di 382 bps rispetto all'esposizione al tasso privo di rischio, rendendola in pratica non efficiente. Nella Tabella 9 vengono riassunti i risultati per le diverse allocazioni.

ANALYTICS E2019 - 2020	MV	Tangenza	EW
E[r]	-1,94%	10,35%	-0,44%
$\sigma$	0,52%	5,2%	1,29%
SR	-7,3462	1,6288	-1,7984
MR	-3,7308	1,9904	-0,3411

Tabella 9. Analytics attese della rete NARX per il periodo 2019-2020

Lo scenario mostra come l'unica strategia efficiente sia quella di tangenza, con uno SR pari 1,6288. Come anticipato, la strategia di minima varianza ne restituisce invece un valore negativo, di cui viene rappresentata l'allocazione in Figura 6.

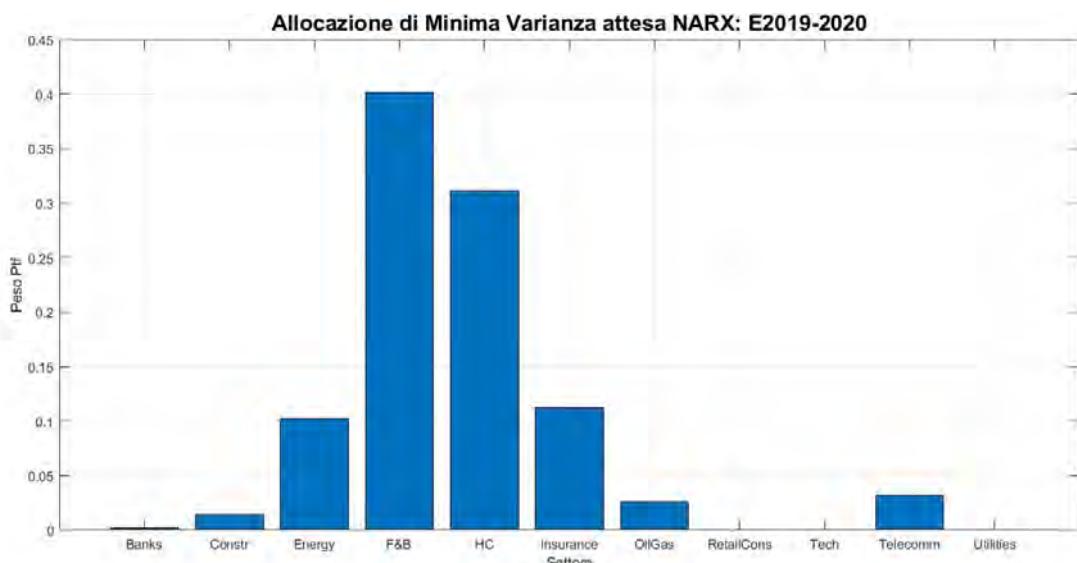


Figura 6. Pesi a MV attesi con rete NARX da Ottobre 2019 ad Ottobre 2020

Lo scenario di tangenza atteso, come rappresentato in Figura 7 presenta un alto grado di concentrazione verso i tre settori migliori, di cui la rete prevede uno scenario rialzista: rispettivamente, i settori bancario, energetico e di oil&gas pesano per il 27,74%, 36,71% ed infine 35,55%.

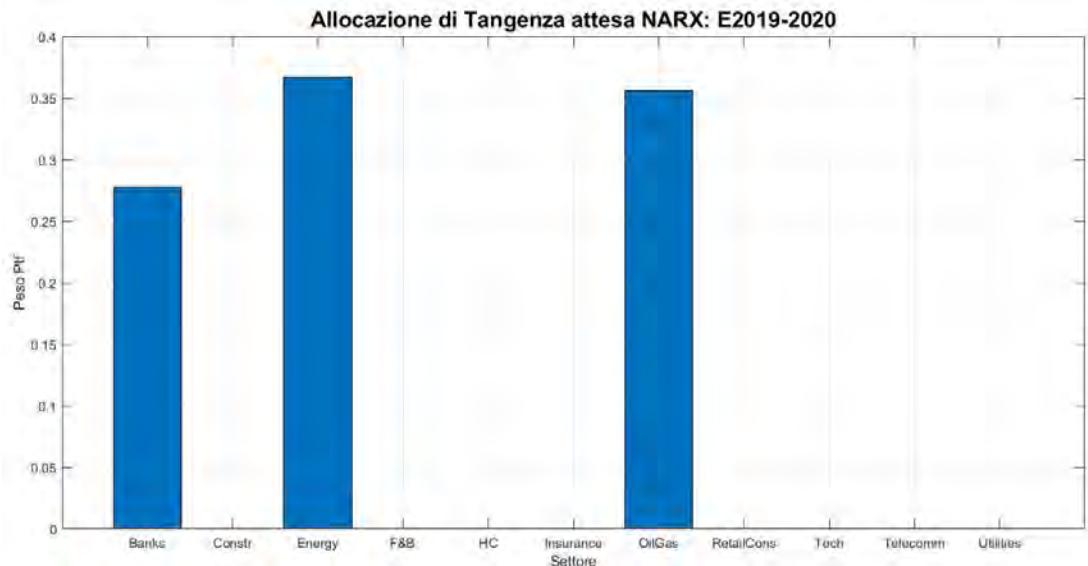


Figura 7. Pesi di tangenza attesi con rete NARX da Ottobre 2019 ad Ottobre 2020

La Figura 8 riassume graficamente l'andamento complessivo previsto, insieme al valore dell'investimento nel titolo risk-free.

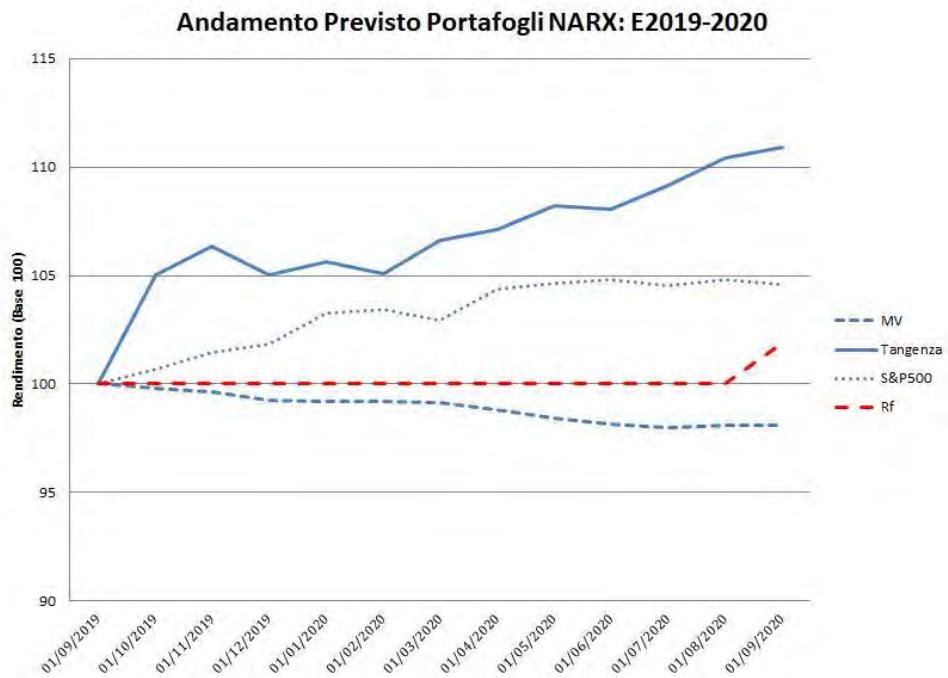


Figura 8. Andamento strategie attese con rete NARX da Ottobre 2019 ad Ottobre 2020

E' infine possibile calcolare il Conditional value-at-risk (altrimenti conosciuto come Expected Shortfall) in chiave prospettica, concentrandosi sui migliori risultati ottenuti, utilizzando la rete con struttura NARX. Vengono definite due frontiere efficienti finali, costruite basando i calcoli sul CVar e seguendo i medesimi vincoli utilizzati in precedenza: investire obbligatoriamente tutta la liquidità disponibile e l'assenza di short selling. Le logiche seguite per la costruzione delle frontiere ricalcano la teoria del portafoglio moderno di Markowitz, utilizzando come variabile di misurazione del rischio il CVar anziché la deviazione standard, scelto come indicatore alternativo poichè capace di fornire una più accurata e ragionevolmente più cautelativa rappresentazione degli scenari avversi. Di fatti, contrariamente al Var, include le variazioni negative che, di norma, verrebbero identificate come outliers e perciò considerabili come non statisticamente rilevanti. Per quanto concerne la definizione delle frontiere, vengono in prima battuta generati 100.000 scenari.

Questi partono dall'assunzione che la distribuzione dei sottostanti sia di tipo Gaussiano, a cui possono dunque essere applicate

le proprietà statistiche delle distribuzioni normali multivariate, ricordando che tale metodologia tiene conto delle deviazioni standard dei rendimenti prospettici degli assets e dei livelli di correlazione esistenti fra di essi. In secondo luogo, è possibile quindi generare le due frontiere efficienti di interesse, definite con 30 portafogli dominanti ciascuna. Vengono rispettivamente rappresentate in Figura 9 ed in Figura 10, con intervalli di confidenza al 95% ed al 99%. I risultati mostrati si intendono calcolati con valori annualizzati.

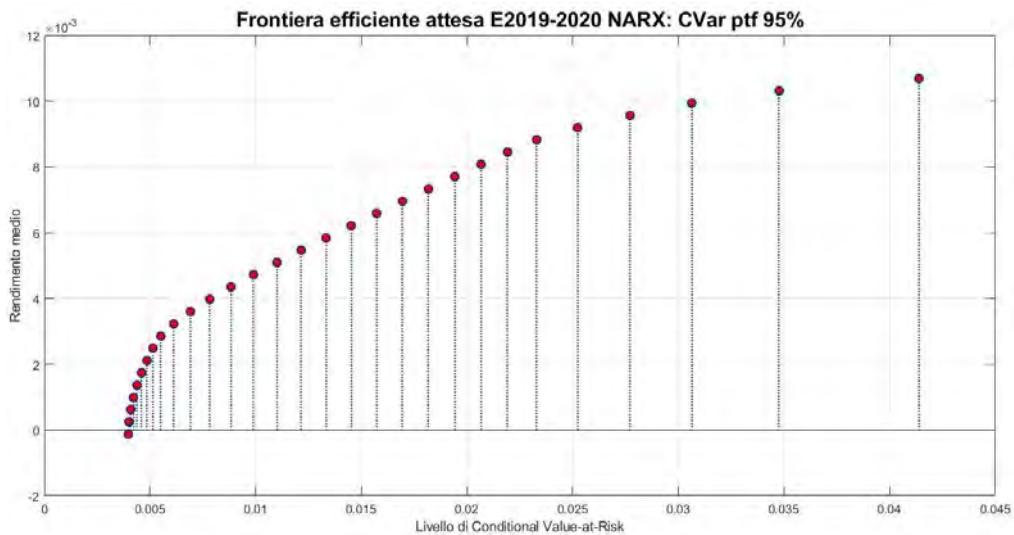


Figura 9. Frontiera efficiente con CVar al 95%

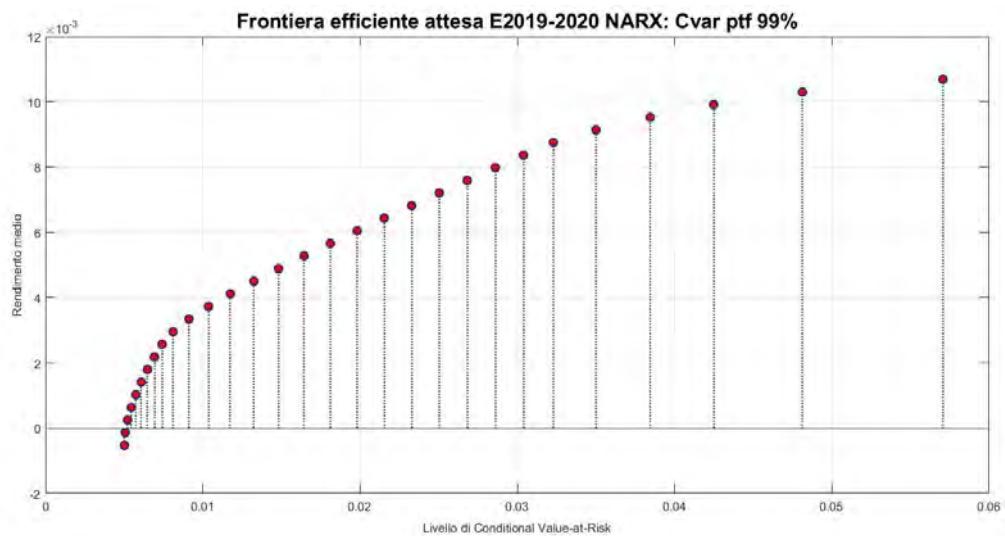


Figura 10. Frontiera efficiente con CVar al 99%

In primo luogo, si evidenzia come il trade-off rischio-rendimento sia, in questo caso, espresso in termini di CVar (sulle ascisse) e di rendimento medio (sulle ordinate) e di come tale divario ragionevolmente aumenti passando dal 95% al 99%. Da un punto di vista di analisi finanziaria, coerentemente con quanto ottenuto in precedenza, la strategia che massimizza il rendimento (e quindi contemporaneamente col valore più alto di CVar) vede la totalità dell'allocazione delle risorse nuovamente nel settore bancario per ambo gli scenari. La strategia che persegue l'obiettivo di minimizzare il CVar, presupponendo rendimento positivo, mostra in entrambi gli scenari un'asset allocation con elevata concentrazione nei settori del food & beverage, dell'healthcare e dell'energia, con pesi rispettivamente pari al 49,8%, 22,8% e 17,85%. In Tabella 10 vengono dunque mostrati i risultati numerici finali ottenuti.

ANALYTICS CVar E2019 – 2020	Minimo CVar [95%]	Max Rendimento [95%]	Minimo CVar [99%]	Max Rendimento [99%]
E[r]	0,025%	1,07%	0,025%	1,07%
CVar	0,4%	4,14%	0,52%	5,71%
MR	0,0625	0,2585	0,0481	0,1874

Tabella 10. Analytics attese CVar

## 5 Conclusioni

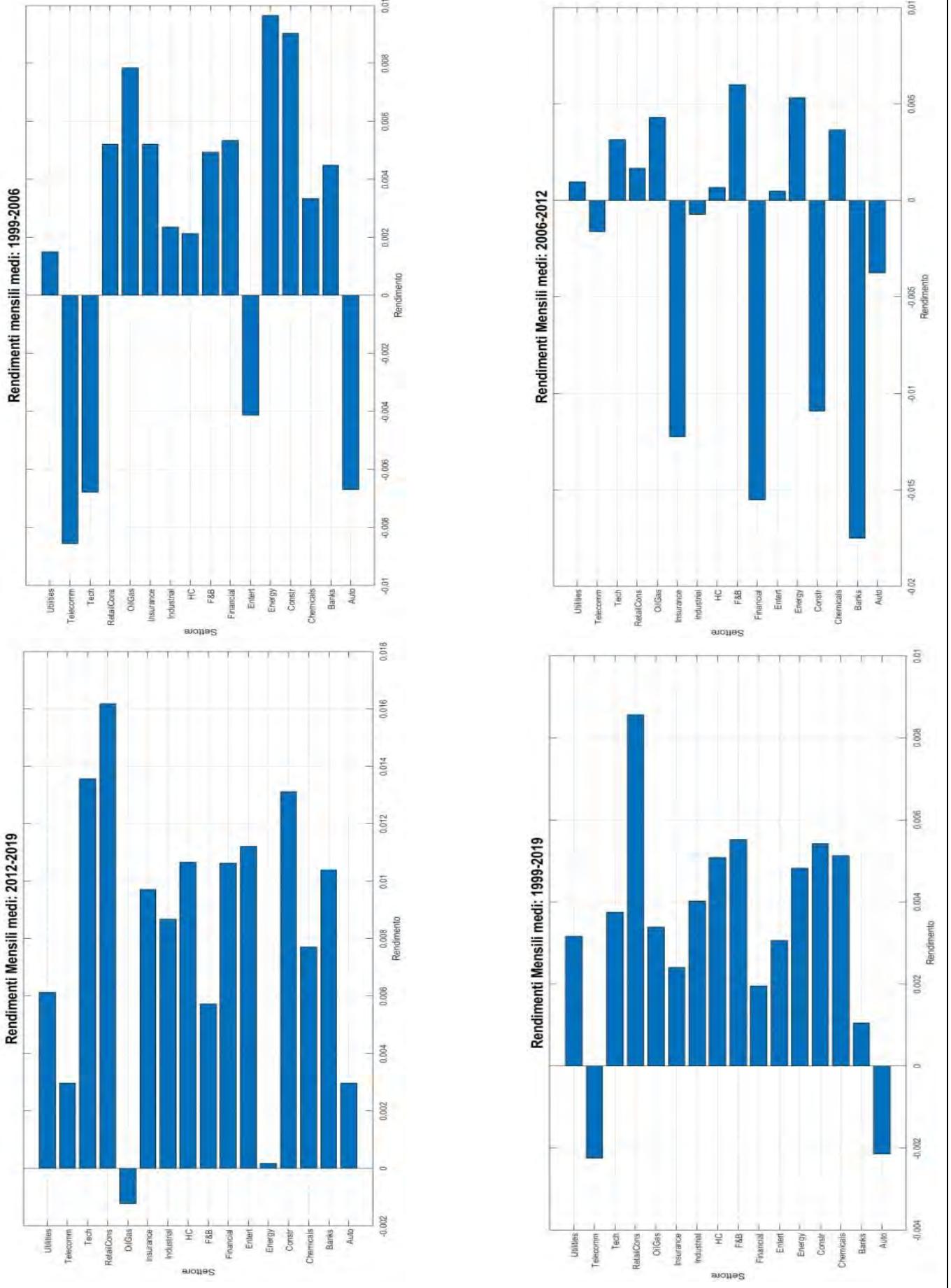
Il paper ha dimostrato come le reti neurali dinamiche involventi variabili esogene (NARX ANN) possano essere utilmente impiegate per il calcolo di misure finanziarie e di rischio di natura forward-looking, individuando una metodologia per disporre di stime affidabili future a partire dai livelli di prezzo degli indici più rappresentativi del mercato americano. Dapprima si è proceduto ad applicare i principi della Modern Portfolio Theory, finalizzati ad individuare quali, tra gli indicatori scelti per rappresentare il mercato, sia risultato più significativo in termini di asset allocation nei periodi temporali considerati finanziariamente rilevanti. Terminata tale fase di vagliatura, sono state addestrate due tipologie di reti neurali dinamiche finalizzate al forecasting: NAR (Nonlinear Auto Regressive) e NARX (Nonlinear Auto Regressive with exogenous variables) network. Lo studio approfondisce le modalità e la misurazione di raggiungimento delle performance del modello tramite indicatori in-sample sul dataset di training ( $R^2$ , MSE ed assenza di autocorrelazione dell'errore al livello di confidenza del 95%) ed out-of-sample tramite back-testing (RMSE, MAE, MSEP). La rete ha mostrato buone capacità di adattamento ai dati di input, fornendo come risposta ottimi risultati in fase di back-testing. Una volta condotti tali test, si è impiegata l'architettura validata di rete per la proiezione dei livelli di prezzo degli indici. Le stime future si sono utilizzate per dare una visione forward-looking delle misure finanziarie e di rischio. A questo punto si è reso possibile procedere oltre, con l'analisi del CVar, considerata una delle misure di rischio più importanti ed impiegate sia in ambito accademico che operativo. I risultati ottenuti hanno mostrato, coerentemente con la realtà empirica, quali fra i settori che sono stati considerati finanziariamente rilevanti in termini di asset allocation, abbiano concorso a minimizzare il livello di rischio in termini prospettici. Si è notata inoltre una variazione nell'allocazione delle risorse rispetto al metodo classico, che vede aumentare la partecipazione nel settore energetico a discapito di quello dell'healthcare, quest'ultimo considerato dal modello più rischioso in termini di drawdown e dunque con impatti diretti sui costi derivanti da maggiori accantonamenti.

Si ritiene interessante, per il proseguo del lavoro, utilizzare all'interno della rete neurale dati macroeconomici in qualità di variabili esogene. Questa integrazione potrebbe ulteriormente migliorare gli indicatori di performance out-of-sample del predittore, soprattutto, come nel caso trattato, in cui si debbano eseguire proiezioni multi step ahead. Inoltre, ai fini di valutazione dei rischi riconducibili ad una determinata asset class o ad una definita asset allocation, l'introduzione di strumenti di *machine learning* come quelli impiegati, unitamente all'analisi di variabili esogene finanziarie e macroeconomiche, può rappresentare per le istituzioni finanziarie un metodo efficace e prospettico di definizione dei livelli di capitale richiesto dal regulator, contribuendo in via preventiva a definire le strategie migliori in caso di eventuali scenari di "stress" finanziario.

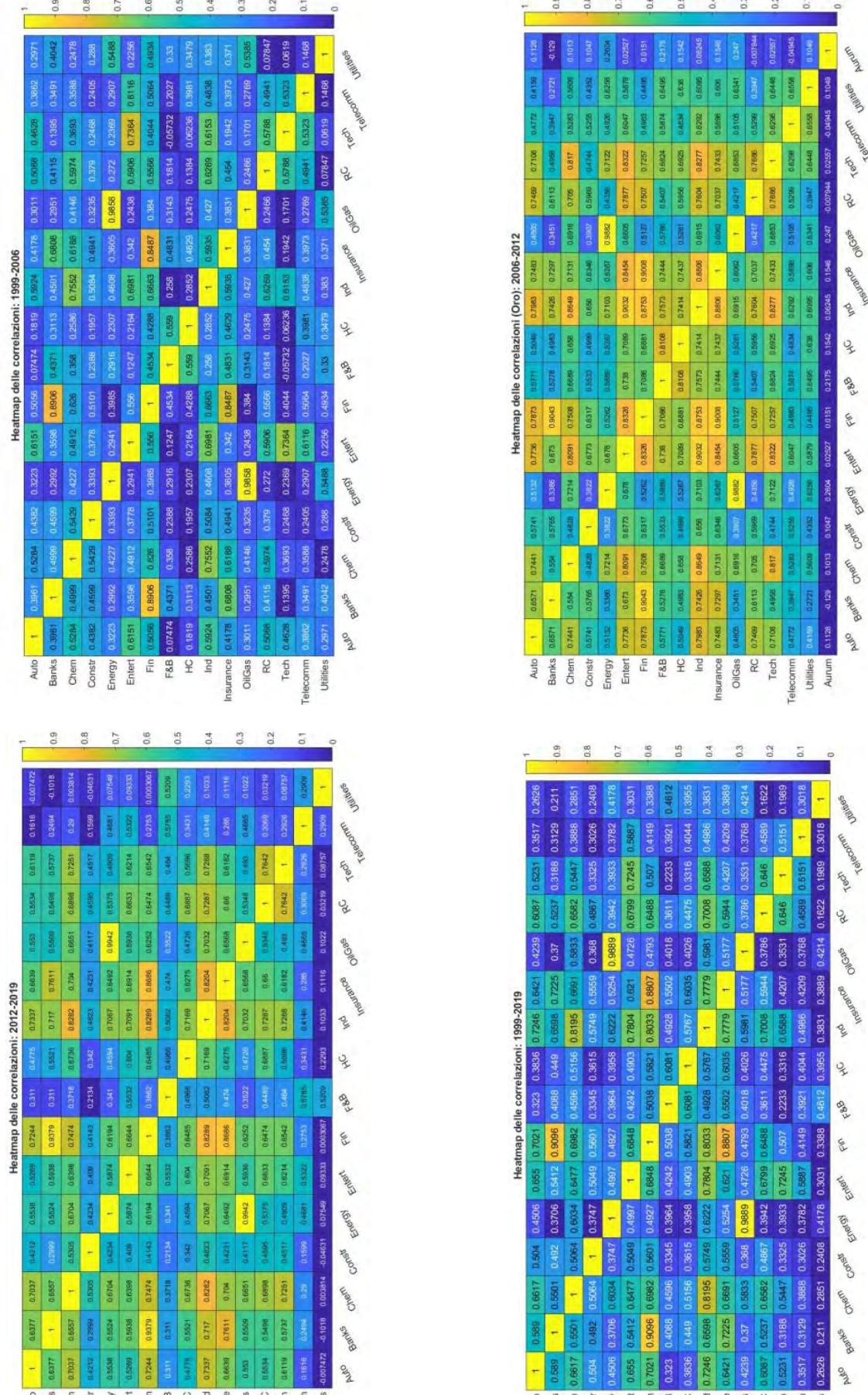
Carlo Decherchi e Pier Giuseppe Giribone

## Bibliografia

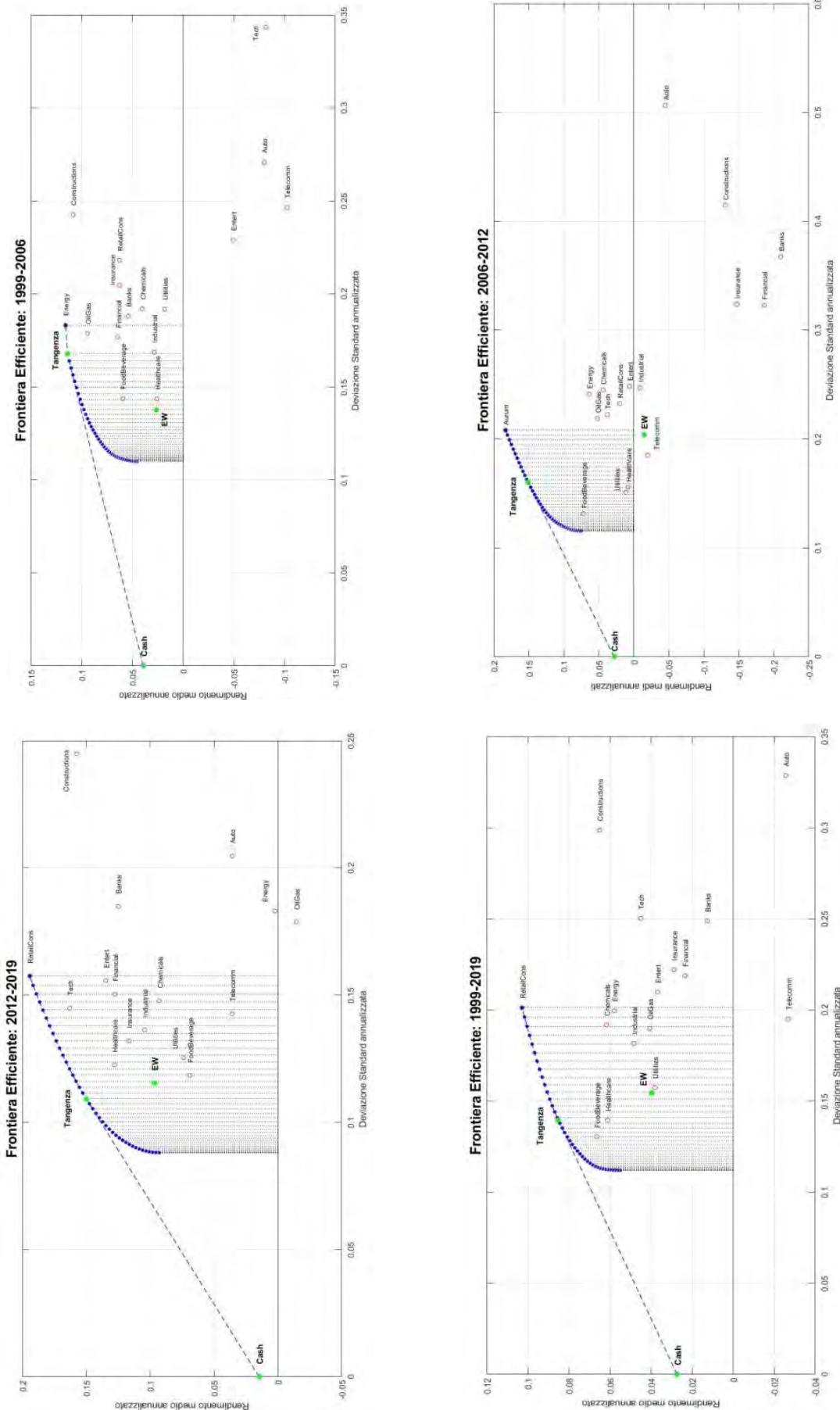
- [1] Arbib, M. A. (2002). *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: MIT Press.
- [2] Beale, M. H. and Hagan, M. T. and Demuth, H. B. (2019). *Neural Network Toolbox User's guide*, MA: Mathworks Press
- [3] Bonini, S. and Caivano, G. and Cerchiello, P. and Giribone, P. G. (2019). Intelligenza Artificiale: l'applicazione di Machine Learning e Predictive Analytics nel Risk Management. XIV Position Paper AIFIRM (Associazione Italiana Financial Industry Risk Managers)
- [4] Caligaris, O. (2007). Le reti neurali. Lettera matematica PRISTEM, 61: 20-28.
- [5] Chang, Z. and Zhang, Y. and Chen, W. (2019). Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. Energy 187: 115804. doi: 10.1016/j.energy.2019.07.134
- [6] Choudhry, M. (2003). *Analysing and Interpreting the Yield Curve*. John Wiley & Sons Inc., pp. 36-59
- [7] De Jesus, O. and Horn, J. M. and Hagan, M. T. (2001). Analysis of Recurrent Network Training and Suggestions for Improvements. Proceedings of the International Joint Conference on Neural Networks, Washington D. C., July 15-19, pp. 2632-2637. doi: 10.1109/IJCNN.2001.938785
- [8] De Leone, R. and Marchitto, E. and Quaranta A.G. (2006). Autoregression and artificial Neural Networks for Financial Market Forecast. Neural Network World 16(2):109-128
- [9] Decherchi, C. (2019). Oltre Markowitz: tecniche dinamiche di forecasting con variabili esogene sul mercato statunitense. Tesi di Laurea Magistrale in Economia ed Istituzioni Finanziarie, Università di Genova. Relatore: prof. Marina Resta, Correlatore: Ing. Pier Giuseppe Giribone
- [10] Giribone, P. G. and Ligato, S. and Penone, F. (2018). Combining robust Dynamic Neural Networks with traditional technical indicators for generating mechanic trading signals. International Journal of Financial Engineering, 5(4): 1850037. doi: <https://doi.org/10.1142/S2424786318500378>
- [11] Horn, J. M. and De Jesus, O. and Hagan, M. T. (2009). Spurious Valleys in the Error Surface of Recurrent Networks: Analysis and Avoidance. IEEE Transactions on Neural Networks, 20(4): 686-700. doi: 10.1109/TNN.2008.2012257
- [12] Markowitz, H. (1992). Portfolio Selection. The Journal of Finance, 7(1): 77-91. doi: 10.2307/2975974
- [13] Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics, 11(2): 431-441
- [14] Resta, M. (2009). Seize the (intra)day: Features selection and rules extraction for tradings on high-frequency data. Neurocomputing, 72: 3413-3427. doi: 10.1016/j.neucom.2008.09.028
- [15] Rojas, R. (1996). *Neural Networks: A systematic introduction*. Springer-Verlag, New York
- [16] Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. The Journal of Finance, 19: 425-442. doi: 10.1111/j.1540-6261.1964.tb02865.x
- [17] Tsay, R. S. (2010). *Analysis of Financial Time Series*. John Wiley & Sons Inc., pp. 199-205



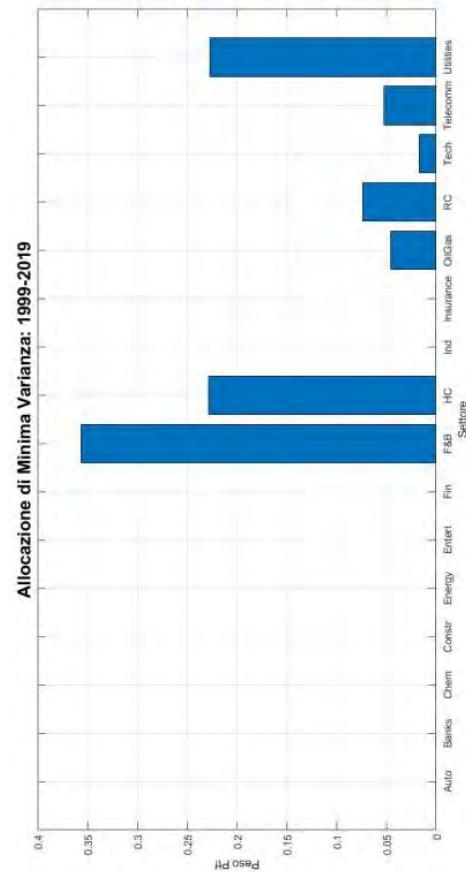
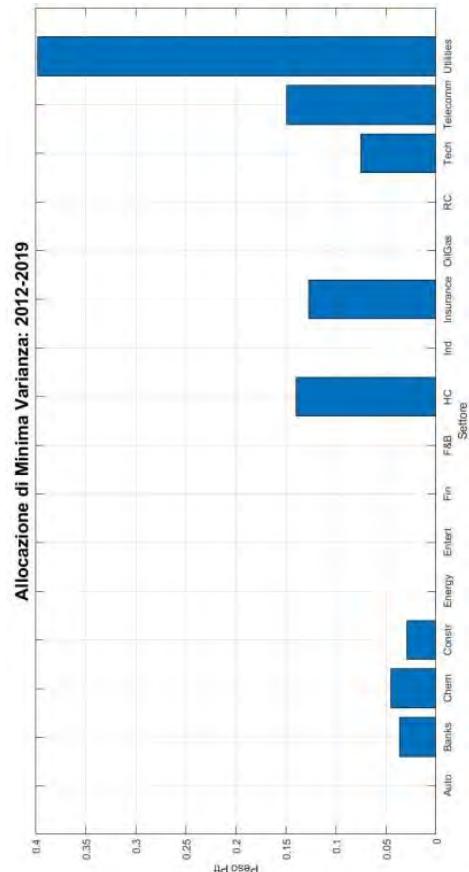
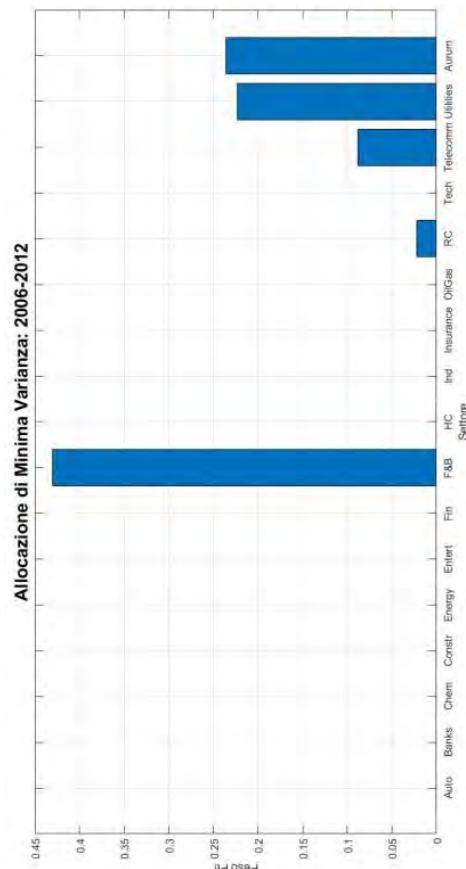
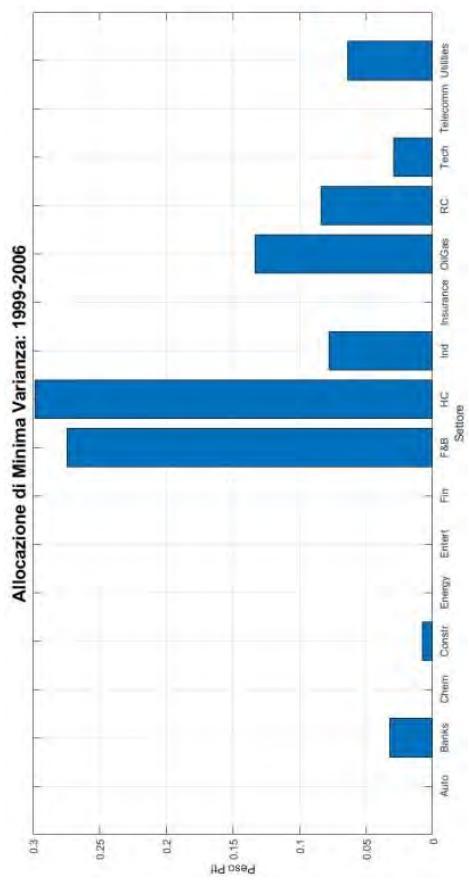
Panel 1. Rendimenti mensili medi



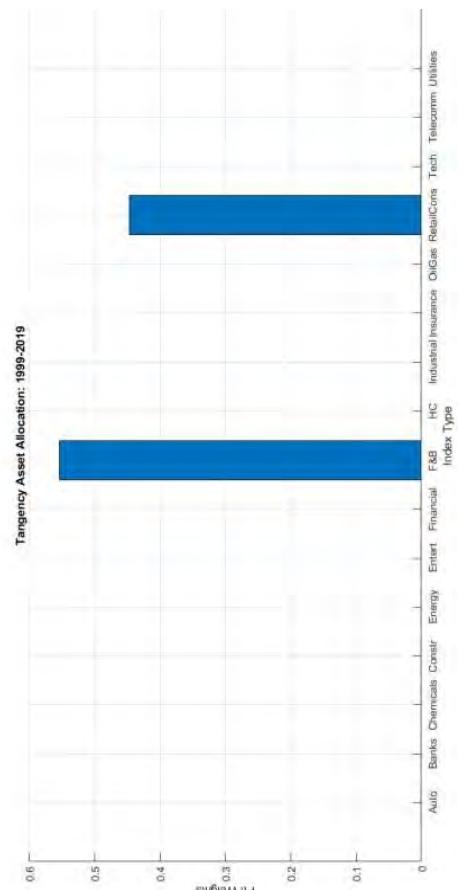
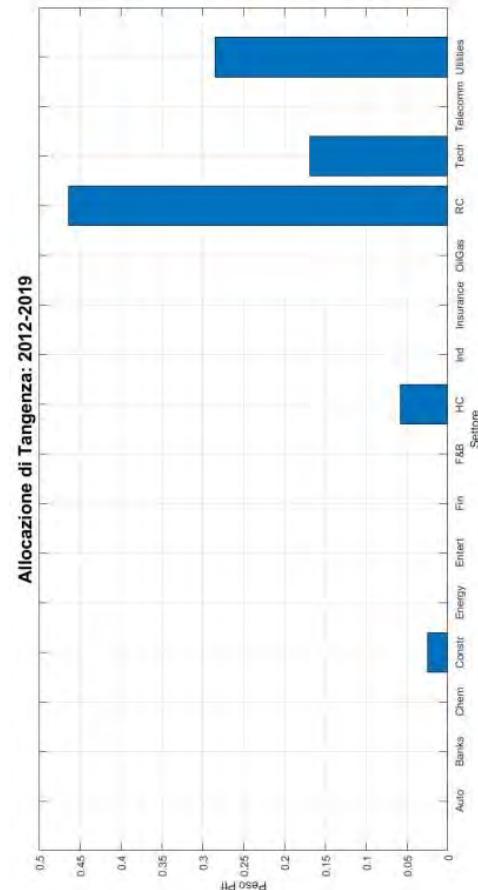
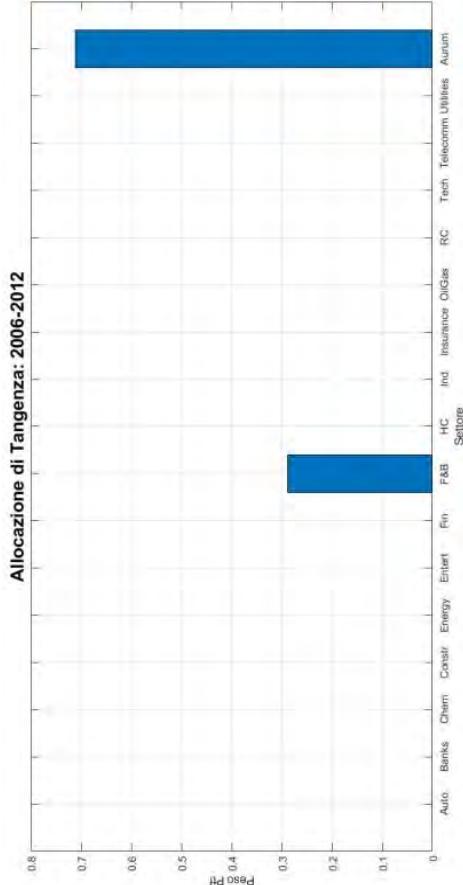
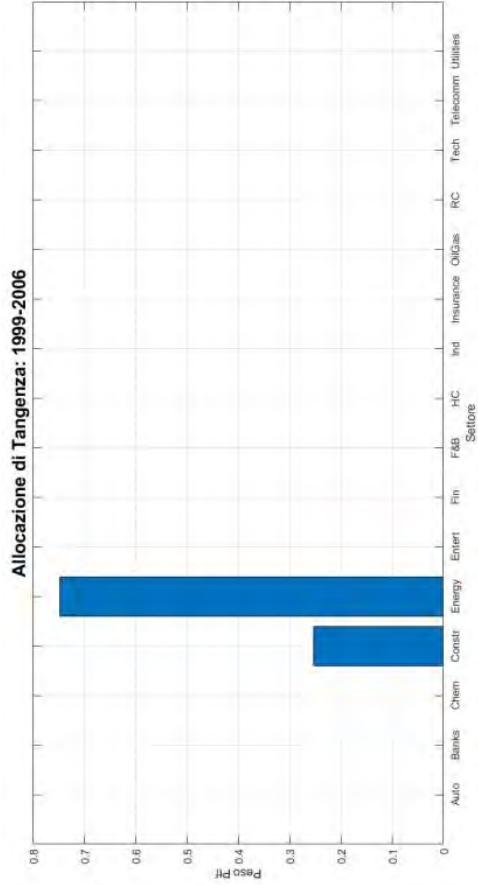
*Panel 2. Heat map delle correlazioni*



Panel 3. Frontiera Efficiente

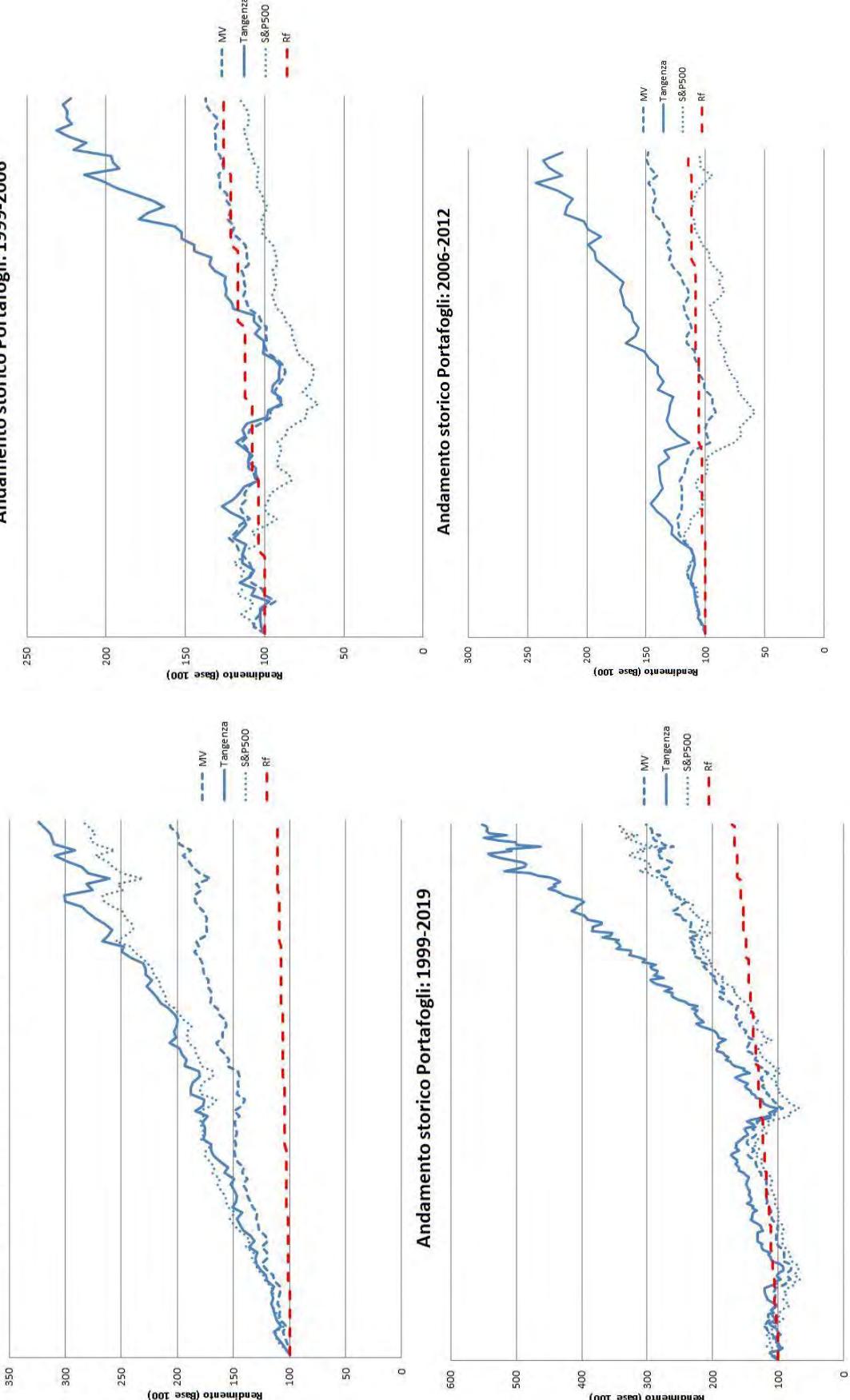


Panel 4. Pesi a Minima Varianza (MV)

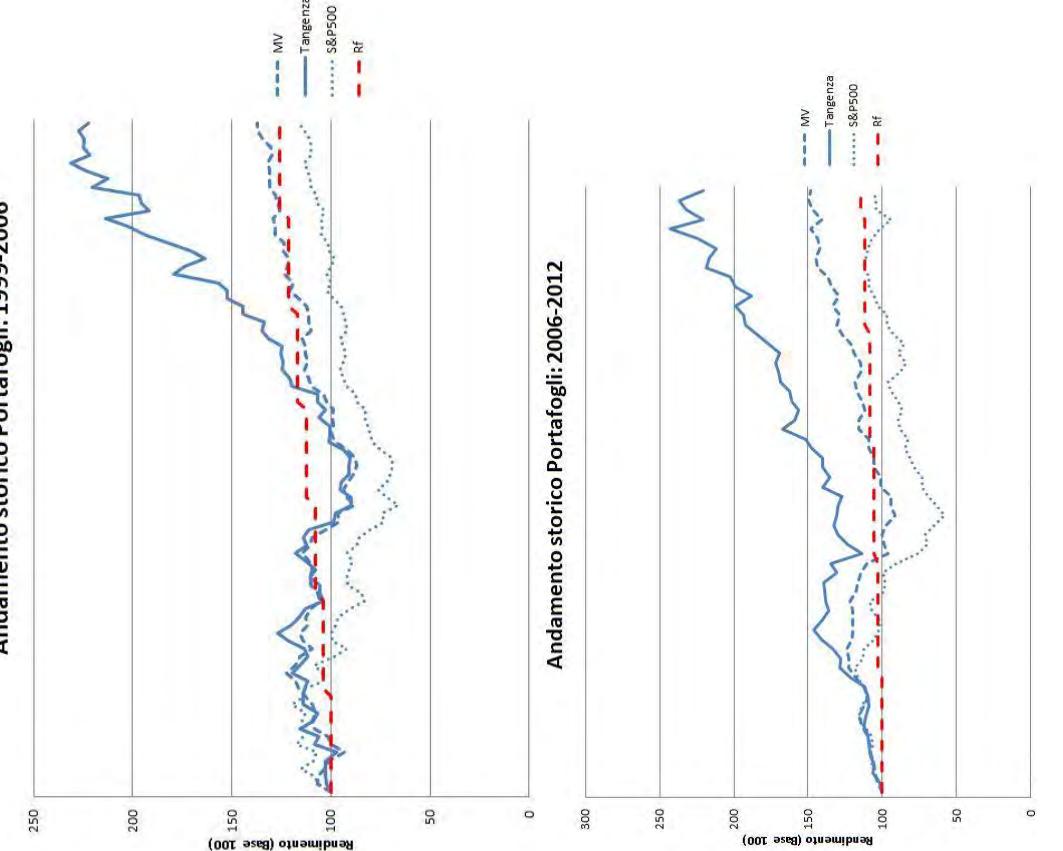


Panel 5. Pesi di tangenza

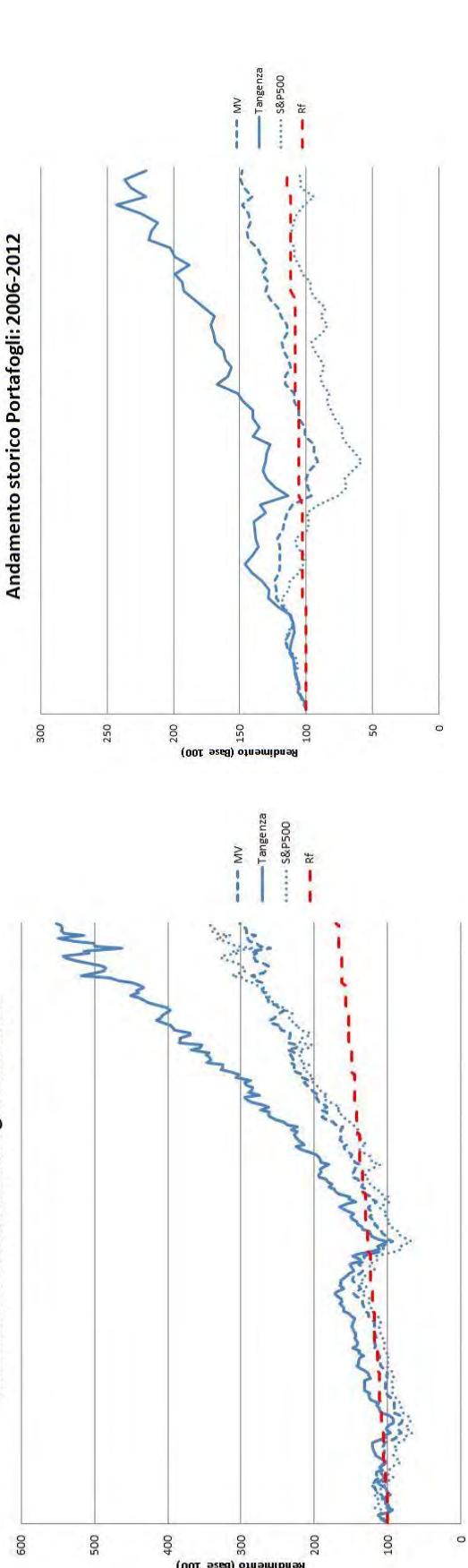
*Andamento storico Portafogli: 2012-2019*



*Andamento storico Portafogli: 1999-2006*



*Andamento storico Portafogli: 1999-2019*



*Panel 6. Andamento delle strategie*

OGGI PARLERO' DELLA  
GESTIONE DEL RISCHIO...



FilVA

## **Risk Management Magazine**

Anno 15, n° 1 Gennaio – Aprile 2020

### **Direttore Responsabile**

Maurizio Vallino

### **Condirettore**

Corrado Meglio

### **Consiglio Scientifico**

Giampaolo Gabbi (Direttore del Consiglio Scientifico),  
Ruggero Bertelli,  
Paola Bongini,  
Anna Bottasso,  
Marina Brogi  
Ottavio Caligaris,  
Simona Cosma,  
Paola Ferretti,  
Andrea Giacomelli,  
Pier Giuseppe Giribone,  
Adele Grassi,  
Valentina Lagasio,  
Duccio Martelli,  
Laura Nieri,  
Pasqualina Porretta,  
Anna Grazia Quaranta,  
Francesco Saita,  
Enzo Scannella,  
Cristiana Schena,  
Giuseppe Torluccio



Rivista dell'Associazione Italiana Financial Industry Risk Managers

### **Comitato di redazione**

Emanuele Diquattro  
Fausto Galmarini  
Igor Gianfrancesco  
Camillo Giliberto  
Rossano Giuppa  
Aldo Letizia  
Paolo Palliola  
Enzo Rocca  
Fabio Salis

**Vignettista:** Silvano Gaggero

### **Proprietà, Redazione e Segreteria:**

Associazione Italiana Financial Industry Risk Managers (AIFIRM), Via Sile 18, 20139 Milano

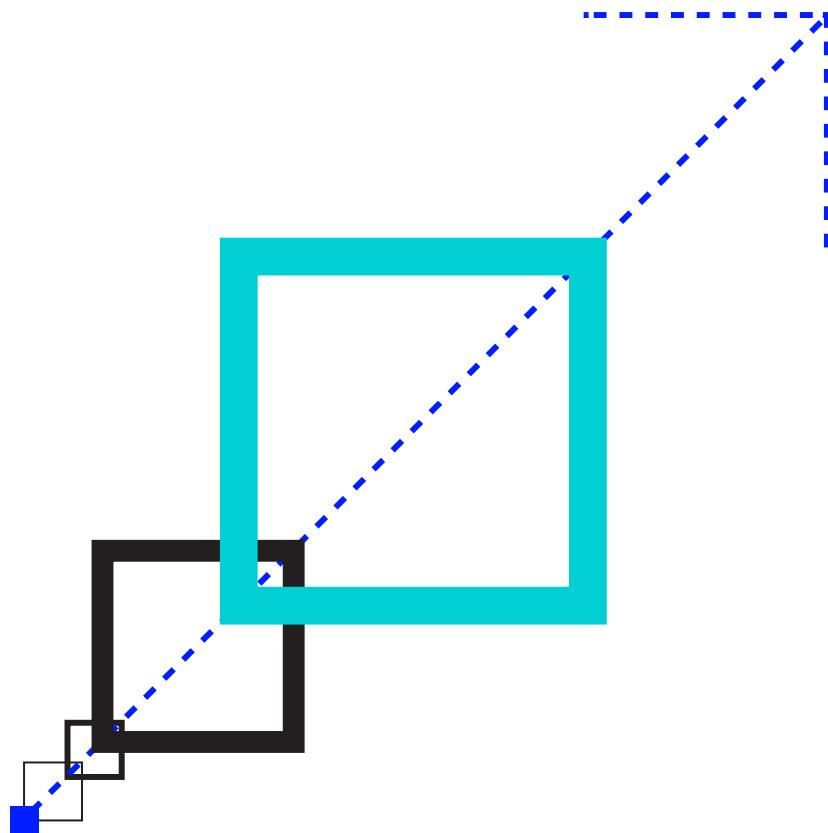
Registrazione del Tribunale di Milano n° 629 del 10/9/2004

E-mail: [segreteria@aifirm.it](mailto:segreteria@aifirm.it); Tel. 389 6946315 Lunedì - Venerdì h.15-17

**Stampa:** Algraphy S.n.c. - Passo Ponte Carrega 62-62r, 16141 Genova

**Le opinioni espresse negli articoli impegnano unicamente la responsabilità dei rispettivi autori**

**SPEDIZIONE IN ABBONAMENTO POSTALE AI SOCI AIFIRM RESIDENTI IN ITALIA, IN REGOLA CON L'ISCRIZIONE**



---

# RELY ON OUR REPUTATION TO PROTECT YOURS

Refinitiv Enhanced Due Diligence reports provide detailed integrity and advanced background checks. With a global, on-the-ground footprint, we leverage local insights and the industry's most trusted financial crime risk screening database.

Our globally recognized experts deliver the trusted perspective to help you make informed business decisions.

[refinitiv.com/edd](https://refinitiv.com/edd)

**REFINITIV®**  
DATA IS JUST  
THE BEGINNING® 