



**UNIVERSITÀ  
DEGLI STUDI  
DI TRIESTE**

---

Dipartimento di Scienze Economiche, Matematiche e Statistiche, Bruno De Finetti

**CORSO DI LAUREA MAGISTRALE IN SCIENZE STATISTICHE E  
ATTUARIALI**

## **I modelli Regime Switching per la valutazione del rischio di estinzione anticipata di massa**

**Un approccio alla costruzione del modello interno per il rischio di estinzione  
anticipata di massa in ottica Solvency II**

***Laureando: Lorenzo Zubin***

***Relatore: Prof. Pietro Millossovich  
Correlatore: Dr. Roberto Bruni***

**ANNO ACCADEMICO 2023 - 2024**

*“Nel suo profondo vidi che s'interna,  
legato con amore in un volume,  
ciò che per l'universo si squaderna:*

*sustanze e accidenti e lor costume  
quasi conflati insieme, per tal modo  
che ciò ch'ì dico è un semplice lume.”*

- Dante Alighieri, *Divina Commedia*,  
Paradiso, canto XXXIII, vv. 85-90

# Indice

<b>Elenco delle Abbreviazioni</b>	<b>1</b>
<b>Introduzione</b>	<b>3</b>
<b>1 La Direttiva Europea <i>Solvency II</i></b>	<b>6</b>
1.1 Una panoramica generale . . . . .	6
1.1.1 Il Primo Pilastro . . . . .	8
1.1.2 Il secondo pilastro: . . . . .	11
1.1.3 Il terzo pilastro: . . . . .	12
1.2 Il Requisito Patrimoniale di Solvibilità . . . . .	12
1.3 Le Riserve Tecniche in ambito Vita . . . . .	17
<b>2 Riscatto e Abbandono</b>	<b>21</b>
2.1 Il diritto di Riscatto e Abbandono e altri diritti nelle polizze vita . . . . .	21
2.2 Il rischio di estinzione anticipata nella Direttiva . . . . .	25
2.3 Le cause del rischio di estinzione e del recesso di massa . . . . .	31
<b>3 La modellizzazione del <i>lapse rate</i> e del <i>mass lapse risk</i></b>	<b>38</b>
3.1 L'approccio ordinario alla modellizzazione del <i>lapse rate</i> . . . . .	38
3.1.1 I metodi della tecnica attuariale Vita per l'analisi del <i>lapse rate</i> . . . . .	38
3.1.2 I metodi della tecnica attuariale Danni per l'analisi del <i>lapse rate</i> . . . . .	40
3.1.3 Gli alberi di Regressione per l'analisi del <i>lapse rate</i> . . . . .	42
3.1.4 L'analisi del <i>lapse rate</i> come serie storica . . . . .	43
3.1.5 Le diverse definizioni del tasso di riscatto . . . . .	47
3.1.6 L'approccio tramite <i>Dynamic Lapse Multiplier</i> per l'analisi del <i>lapse rate</i> . . . . .	51
3.2 Il problema del <i>mass lapse</i> e le proposte in letteratura . . . . .	52
3.2.1 Biagini, Huber, Jaspersen e Mazzon 2020 . . . . .	55
3.2.2 Loisel e Milhaud 2011 . . . . .	55
3.2.3 Barsotti, Milhaud e Salhi 2016 . . . . .	56
3.2.4 Cheng, Hilpert, Lavasani e Schaefer 2023 . . . . .	57
<b>4 I modelli <i>Regime Switching</i></b>	<b>58</b>
4.1 I Modelli TAR . . . . .	60
4.2 I Modelli STAR . . . . .	66
4.3 Il " <i>Modelling Cycle</i> " per i modelli STAR . . . . .	71
4.4 Un modello con <i>trend</i> deterministico non lineare . . . . .	72
4.5 Le critiche ai modelli . . . . .	76

<b>5</b>	<b>Analisi empirica</b>	<b>78</b>
5.1	Il campione . . . . .	78
5.2	Le Variabili Esterne . . . . .	88
5.3	La Stima dei Modelli . . . . .	91
5.3.1	La stima del modello AR . . . . .	92
5.3.2	La stima del modello TAR . . . . .	96
5.3.3	Stima dei Modelli LSTAR . . . . .	104
5.3.4	Stima del modello con <i>trend</i> deterministico . . . . .	108
5.3.5	La valutazione del <i>mass lapse risk</i> . . . . .	113
	<b>Conclusioni</b>	<b>117</b>
	<b>Appendici</b>	<b>119</b>
5.4	La Stazionarietà . . . . .	119
5.5	I Modelli AR . . . . .	121
5.6	Il massimo della curva Gamma del modello a <i>trend</i> deterministico . . . . .	123
	<b>Ringraziamenti</b>	<b>125</b>
	<b>Bibliografia</b>	<b>126</b>

## Elenco delle Abbreviazioni

- **ACF**: *Autocorrelation Function*;
- **ADL**: modello *Autoregressive Distributed Lag*;
- **ANIA**: Associazione nazionale fra le imprese assicuratrici;
- **AR**: modello Autoregressivo;
- **ARIMA**: modello Autoregressivo a Media Mobile Integrato;
- **ARMA**: modello Autoregressivo a Media Mobile;
- **BEL**: *Best Estimate of Liabilities*;
- **BSCR**: *Basic Solvency Capital Requirement*;
- **BTP**: Buoni del Tesoro Poliennali;
- **CEIOPS**: *Committee of European Insurance and Occupational Pensions Supervisors*;
- **DPHB**: *Dynamic PolicyHolder Behaviour*;
- **EIOPA**: *European Insurance and Occupational Pensions Authority*;
- **ERM**: *Enterprise Risk Management*;
- **ESTAR**: modello *Exponential Smooth Transition Autoregressive*;
- **EVT**: *Extreme Value Theory*;
- **FdR**: Funzione di Ripartizione;
- **GLM**: *Generalized Linear Models*;
- **HRG**: *Homogeneous Risk Group*;
- **IASB**: *International Accounting Standards Board*;
- **IFRS**: *International Financial Reporting Standards*;
- **IVASS**: Istituto per la Vigilanza sul Sistema Assicurativo;
- **LM**: *Linear Models*;
- **LOB**: *Line of Business*;
- **LSTAR**: modello *Logistic Smooth Transition Autoregressive*;
- **MCR**: *Minimum Capital Requirement*;
- **NAV**: *Net Asset Value*;
- **NLS**: stimatore *Non linear Least Squares*;
- **OLS**: stimatore *Ordinary Least Squares*;
- **PACF**: *Partial Autocorrelation Function*;
- **PHB**: *PolicyHolder Behaviour*;
- **POT**: metodo *Peaks Over Threshold*;
- **QIS**: *Quantitative Impact Studies*;
- **RM**: *Risk Margin*;
- **RSM**: *Regime Switching Models*;
- **RT**: *Riserve Tecniche*;
- **SCR**: *Solvency Capital Requirement*;
- **SESTAR**: modello *Self Exciting Smooth Transition Autoregressive*;

- **SETAR:** modello *Self Exciting Threshold Autoregressive*;
- **STAR:** modello *Smooth Transition Autoregressive*;
- **TAR:** modello *Threshold Autoregressive*;
- **TP:** *Technical Provisions*;
- **TVaR:** *Tail Value at Risk*;
- **USP:** *Undertaking Specific Parameters*;
- **VaR:** *Value at Risk*;

# Introduzione

Il 6 febbraio del 2023 si verifica un evento mai osservato nel panorama assicurativo italiano, e solamente una volta a livello europeo: l'IVASS - l'Istituto di Vigilanza sul Sistema Assicurativo Italiano - sospende il diritto di riscatto a 413 mila polizze, vendute dalla compagnia assicurativa “Eurovita S.p.a.”, operante nel ramo Vita.<sup>1</sup> Il motivo: l'Istituto reputa l'impresa esposta ad un elevato rischio di estinzione anticipata e non in grado di sostenere tale rischio in base alle sue attuali condizioni economiche.

Eurovita era già da molto sotto sorveglianza dell'Autorità Vigilante, la quale aveva suggerito all'azionista di maggioranza, il fondo inglese *Cinven*, già a metà del 2022, di fare un apporto di capitale all'impresa di almeno 150 milioni per risanare una situazione patrimoniale molto debilitata. La compagnia era arrivata infatti ad un *solvency ratio* pari al 117% a fine giugno, contro la media nazionale del 230%, a causa di una cattiva gestione degli investimenti.

La situazione per Eurovita è andata peggiorando da giugno 2022 fino al fatidico 6 febbraio. L'innesto di capitale richiesto è andato aumentando, e ciò a causa dell'aumento repentino dei tassi d'interesse di mercato, per effetto delle decisioni della Banca Centrale Europea. L'aumento dei tassi di mercato, infatti, ha ridotto notevolmente il valore degli investimenti fatti dalla compagnia assicurativa. Se questa si trovasse nella situazione di dover affrontare un ingente esborso di capitale, si potrebbe veder costretta a svendere i propri *asset* e a realizzare tutte le minusvalenze, con la possibile conseguenza di un default. Sfortunatamente per Eurovita, l'aumento dei tassi è quasi sempre correlato - e qui non è accaduto diversamente - con un aumento dei riscatti: quando i rendimenti delle polizze sono inferiori rispetto ai rendimenti dei mercati, gli assicurati sono incentivati a chiudere il contratto per ricevere il valore di riscatto e a reinvestirlo in titoli più remunerativi. Conseguentemente, la compagnia si trova ad affrontare un aumento dei riscatti e quindi comincia a subire grosse perdite, e le prospettive di maggiori perdite si fanno sempre più ingenti.

A inizio 2023, però, poco prima dell'intervento dell'IVASS, Eurovita è ancora solubile, riesce a pagare tutti i riscatti, ciò che però non riesce a sopportare è l'aumento di capitale richiesto per far fronte al rischio di estinzione anticipata, e in particolare al rischio di estinzione anticipata di massa, o “*mass lapse risk*”, la cui valutazione dipende da quanto prescritto dalla normativa vigente, ovvero la Direttiva *Solvency II*. Perciò, Per quanto tecnicamente in positivo, il requisito di capitale imposto dal Legislatore per sopportare il *mass lapse risk* supera di molto il capitale disponibile.

La parabola di Eurovita si concluderà il 27 ottobre del 2023 con il provvedimento dell'IVASS che predisporrà la liquidazione coatta amministrativa, dopo che Civen sceglie di non apportare il capitale richiesto. I contratti verranno trasferiti ad una so-

---

<sup>1</sup>Per una trattazione più completa sul caso Eurovita, si rimanda al sito: <https://soldiexpert.com/moneyreport/polizze-eurovita-blocco-riscatti-rischi-sottoscrittori/136049/>.

cietà costruita ad hoc, la “*Cronos Vita*”, ripartita tra importanti compagnie assicurative italiane, per poter garantire agli assicurati la continuazione dei contratti.

Il caso, come detto, è stato singolare, ma ha avuto una risonanza su tutte le compagnie italiane; ha messo in luce, come mai prima, molti degli aspetti critici della valutazione del rischio di riscatto e del riscatto in massa, e ha fatto sorgere molti interrogativi e preoccupazioni in merito ai metodi utilizzati dalle assicurazioni per la valutazione di tali rischi.

Rifacendosi a queste preoccupazioni, lo scopo della presente Tesi sarà quello di svolgere un’indagine approfondita su questi importanti interrogativi e di fornire delle possibili risposte e proporre delle soluzioni, una in particolare al di fuori dei modelli attualmente proposti. In sintesi, le domande che verranno affrontate saranno le seguenti: *Da che cosa dipende la scelta di riscattare di un assicurato?*, *Quali sono le condizioni che possono portare ad un riscatto di massa?*, *L’approccio previsto da Solvency II per la valutazione del rischio di massa è valido?*, *È possibile definire altri modelli per la valutazione di tale rischio?* e *Questi modelli possono fornire stime più valide per il rischio in esame rispetto a quelle ottenute tramite l’approccio fornito dal Legislatore?*

Nello specifico, il tema centrale della presente Tesi sarà il “rischio di estinzione di massa”, o “*mass lapse risk*”, così come inquadrato all’interno della Direttiva europea “*Solvency II*”. L’intento sarà quello di definire il rischio, descrivere le variabili da cui dipende, osservare quali sono le criticità che lo rendono difficile da modellizzare, fornire un possibile approccio alla sua valutazione, verificare questo possibile approccio su di un campione reale e confrontare i risultati con quelli forniti dalla Direttiva.

Più nel dettaglio, la Tesi si comporrà di cinque capitoli principali, suddivisi in paragrafi e sottoparagrafi, più un capitolo finale conclusivo, a cui seguiranno le Appendici e le Fonti Bibliografiche.

Il *primo capitolo* sarà dedicato alla presentazione della summenzionata Direttiva europea *Solvency II*. Verrà fatta un’esposizione sintetica e mirata alle sole parti di diretto o indiretto interesse per l’argomento in questione. Dopo una panoramica generale sullo sviluppo storico, volta a sottolineare le cause che hanno reso necessaria la sua emanazione, ed una breve descrizione del “sistema a tre Pilastri”, si passerà poi, più nel dettaglio, ad approfondire il primo fra questi, riguardante i requisiti quantitativi richiesti alle Compagnie Assicurative, riferiti al Capitale Proprio e alle Passività. Si parlerà pertanto del “Capitale di Solvibilità”, delle “Riserve Tecniche” e del “Margine per il Rischio”. L’argomento sarà trattato in chiave attuariale e non giuridica, mettendo a risalto gli aspetti inerenti la materia.

Con il *secondo capitolo* si andrà a presentare il tema cuore della Tesi: il “*mass lapse risk*”, così come definito dalla Direttiva. Si partirà da un’introduzione alla problematica generale dei diritti contrattuali degli assicurati e del loro “costo”, soprattutto in termini di rischio e di allocazione del capitale. Successivamente, si passerà all’inquadramento in ottica *Solvency* del fenomeno, dalla definizione di “*lapse*” e di “*lapse risk*”, alla valutazione tramite formula standard di questa e all’impatto sulle imprese derivante dal suo utilizzo. In fine, si andrà a fare una disamina approfondita della attuale teoria



sul comportamento dell'assicurato recedente e di quali sono le variabili che condizionano il *lapse rate* e in particolar modo il *mass lapse*.

Il *terzo capitolo* apre la sezione fondamentale del presente lavoro, ovvero quella riguardante la modellizzazione del rischio di estinzione di massa. Il terzo in particolare si occuperà, nella prima parte, di presentare i diversi approcci correntemente usati per la modellizzazione del *lapse rate*, in particolare del “*Dynamic Policyholder Behaviour*”. Nella seconda parte, si passerà invece ad esibire i diversi aspetti distintivi dell'evento di massa che lo rendono così particolarmente difficile da valutare tramite gli approcci descritti nella prima parte. Conseguentemente verranno presentati, brevemente, quattro articoli incentrati sul tema, proponenti diversi metodi per la risoluzione delle criticità osservate.

Con il *quarto capitolo* si andrà invece a definire un ulteriore approccio possibile per la modellizzazione del rischio in esame, basato sull'utilizzo dei “*Regime Switching Model*”. Verrà presentata l'idea alla base che accomuna questa famiglia di modelli e come essi possano rappresentare il fenomeno del *mass lapse*. Via via di seguito saranno definiti i tre modelli principali che verranno utilizzati per la valutazione del rischio: i modelli *Threshold Autoregressive*, i modelli *Self Exciting Threshold Autoregressive* e i modelli *Smooth Transition Autoregressive*. Il capitolo si concluderà con la presentazione di un ulteriore approccio possibile per la modellizzazione basato sui modelli a *trend* deterministico e con la discussione sui possibili problemi relativi all'applicazione di tale famiglia al problema in esame.

Il capitolo finale, il *quinto capitolo*, concluderà la sezione sulla modellizzazione fornendo una valutazione empirica del nuovo approccio proposto. L'analisi statistica verrà condotta su un campione, proveniente da una compagnia assicurativa italiana, formato da tre serie di tassi di riscatto osservati. Tale campione, a causa della sua numerosità esigua, verrà aumentato tramite informazioni esterne riguardanti il mercato assicurativo italiano. Il fine del capitolo sarà quello di verificare la bontà dei modelli nel rappresentare il fenomeno del rischio di estinzione di massa. Inoltre, tramite le stime prodotte, si andrà ad ottenere le valutazioni del rischio in esame, in particolare stimandone l'impatto per un fissato livello di probabilità. Tali stime verranno poi confrontate con i valori specificati dalla Direttiva Europea

A conclusione della ricerca seguirà, come già detto, un breve capitolo riassuntivo del lavoro svolto, il quale presenterà anche le possibili critiche e migliorie che potrebbero essere apportate alla presente ricerca e ai modelli scelti.

# 1 La Direttiva Europea *Solvency II*

## 1.1 Una panoramica generale

Il termine “*Solvency II*” è il soprannome con cui è comunemente nota la Direttiva europea 2009/138/CE, costituente l’attuale corpus delle leggi comunitarie in materia di accesso ed esercizio dell’attività di assicurazione e riassicurazione. Essa fu definitivamente approvata dal Consiglio Europeo il 25 novembre del 2009 e divenne operativa l’uno gennaio del 2016, a seguito di un esteso processo di studi quantitativi e numerose modifiche e interruzioni, in particolare nel biennio 2010-2011 a causa della crisi dei debiti sovrani e successivamente nel 2014 a seguito dell’introduzione di un’altra direttiva, la *Omnibus II*, avente per tema la tutela dei consumatori<sup>2</sup>. Viene denominata *Solvency II* in quanto sostituisce le precedenti disposizioni comunitarie in materia assicurativa, le quali erano raccolte in una Direttiva allora detta “*Solvency*”, ora “*Solvency Primo*”.

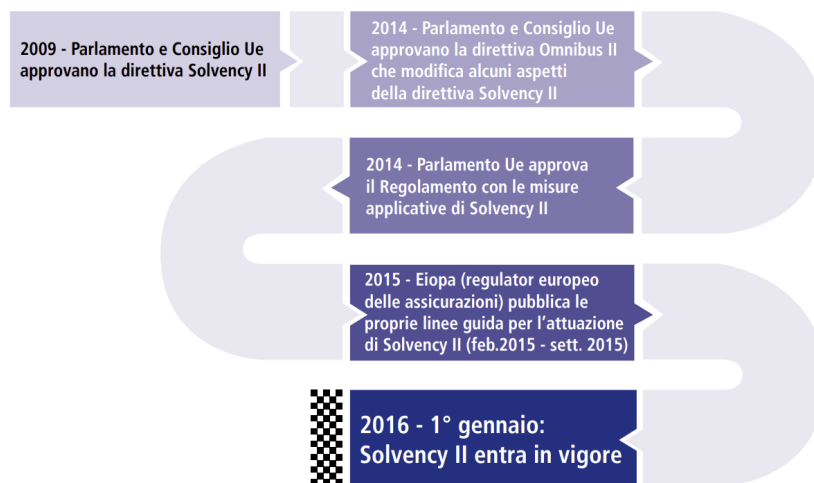


Figura 1: Le tappe rilevanti del percorso storico fatto dalla Direttiva. Fonte: *SOLVENCY II La Nuova Regolamentazione Prudenziale Del Settore Assicurativo: Una Guida Semplificata* (2016).

L’oggetto del teso normativo è riassunto in tre punti chiave: l’accesso e l’esercizio dell’attività di assicurazione e riassicurazione internamente alla Comunità Europea, la vigilanza sui gruppi e sul settore assicurativo e il risanamento e la liquidazione delle

<sup>2</sup>Per una disamina più approfondita sulla storia e sul tema si rimanda al sito dell’IVASS: <https://www.ivass.it/normativa/focus/solvency/index.html>.

imprese assicurative, e, come suggerito dal nome stesso, ha come obbiettivo la garanzia della “solvibilità” delle compagnie. Con questo termine si intende la capacità dell’impresa di adempiere agli obblighi assunti nei confronti di tutte le parti interessate, con specifico riguardo per gli assicurati. La posizione debitoria nei confronti di questi è però tipicamente dipendente da eventi di natura aleatoria, ha quindi un carattere aleatorio. Per questo motivo la nuova Direttiva intende la solvibilità solo in termini aleatori e fissa la garanzia di solvibilità per un determinato livello di probabilità. Esso costituisce il principale punto di innovazione rispetto alla precedente direttiva, la quale si caratterizzava per una debole o assente valutazione dei rischi che sorgono da questa dinamica aleatoria e che gravano su questo particolare settore. Tale mancanza ha lasciato impreparato il mercato assicurativo di fronte alla crisi del 2008, che ha portato al crollo di alcune compagnie assicurative europee.

Da questo approccio cosiddetto “*Risk-based*” di *Solvency II* nasce il passaggio dall’uso di soli modelli “deterministici”, che forniscono singoli valori, all’affiancamento a questi di modelli “stocastici”, cioè che producono distribuzioni ed in particolare stime per la variabilità e quindi per l’errore nelle previsioni. Ovviamente il discorso sulla solvibilità e la valutazione del rischio non si è esaurito nella sola aggiunta di requisiti “quantitativi”, ma anche di requisiti “qualitativi” e “informativi”. A sostenere infatti la solvibilità dell’“edificio-impresa”, il Legislatore ha definito tre “*pilastr*”, che potessero fare da fondamenta. Questi sono:

- Primo pilastro: riferito ai requisiti quantitativi o ai “*Minimum Financial Requirements*”;
- Secondo pilastro: riferito ai requisiti qualitativi o di “*Governance, Internal Control e Risk management*”<sup>3</sup>;
- Terzo pilastro: riferito ai requisiti informativi per l’autorità di vigilanza e per i mercati o di “*Supervisory Reporting e Public Disclosure*”<sup>4</sup>.

Il sistema messo in atto va quindi ben al di là della semplice analisi patrimoniale, ma si preoccupa di tutta quanta l’attività che gli assicuratori svolgono, sia nei suoi processi interni che esterni. È importante sottolineare però che la direttiva non presenta contenuto di natura contabile, o meglio non definisce uno standard di scrittura del bilancio e di redazione della contabilità, ma si limita, in questo ambito, a stabilire i criteri di valutazione delle voci patrimoniali, come sarà detto più avanti. Diverso è il caso degli IFRS 17, che sono la raccolta degli standard contabili internazionali per la scrittura dei bilanci consolidati per le imprese che svolgano l’attività assicurativa e riassicurativa, emanato dall’*International Accounting Standard Board*, o IASB in breve, organo internazionale che si occupa appunto dell’uniformazione degli stati patrimoniali e dei conti economici delle imprese operanti su stati aderenti allo IASB<sup>5</sup>.

---

<sup>3</sup>Dal sito dell’IVASS: <https://www.ivass.it/normativa/focus/solvency/index.html>.

<sup>4</sup>Sempre dal sito dell’IVASS, in generale non c’è una nomenclatura standardizzata dei tre pilastri.

<sup>5</sup>Riguardo agli IFRS 17 si rimanda al sito degli IFRS: <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-17-insurance-contracts/>.

La solvibilità è di fatto il principio cardine attorno a cui ruota la Direttiva. Per fornire un quadro più completo, però, è necessario menzionare anche altri principi fondamentali, non pienamente realizzati da *Solvency I*: in primo luogo il “*principio di armonizzazione*”, alla base di tutte le più recenti direttive comunitarie, il quale qui si realizza attraverso la conformazione delle leggi in materia assicurativa e riassicurativa nei diversi paesi membri, rendendo più semplice alle imprese operare su più stati membri e per garantire pari trattamento giuridico pur avendo sede legale diversa, così da poter favorire la concorrenza nel mercato unico. Altro principio, in sintonia con quanto dettato dallo IASB, è la valutazione “*Market Consistent*” delle Attività e Passività e la loro conseguente iscrizione nello Stato Patrimoniale al “*Fair Value*”<sup>6</sup>. Questo principio prevede che le valutazioni delle Attività e delle Passività dell’impresa siano consistenti con i mercati, ovverosia: il valore contabilizzato di un certo bene o servizio deve equivalere al valore di mercato dello stesso o di un suo simile. Ciò non è in generale immediato e risulta essere un punto ancor più critico per il caso assicurativo, poiché le Passività di una compagnia assicurativa non trovano di norma un mercato di riferimento. La Direttiva si è perciò preoccupata di stabilire delle regole e delle metodologie per poter ottenere valutazioni coerenti con il principio; il punto sarà chiarito in seguito. In fine, è opportuno richiamare il “*principio di prudenza*”, già racchiuso nel concetto di solvibilità, “*il principio di trasparenza*”, “*di chiarezza*” e “*di comparabilità*”, propri anche questi dei Principi Contabili Internazionali.

Tornando al tema del *sistema a tre pilastri*, si possono ora meglio definire nelle loro diverse componenti:

### 1.1.1 Il Primo Pilastro

Esso definisce due presidi patrimoniali: il primo è il Capitale di Solvibilità o il “*Solvency Capital requirement*”, solitamente abbreviato con SCR, a fianco del quale sono posti i criteri per la determinazione dei “*Fondi Propri Ammissibili*” per la costituzione di tale requisito di capitale; il secondo sono le *Riserve Tecniche* e le relative “*Attività Ammissibili*” a copertura di tali riserve.

L’SCR è il livello minimo di Capitale Proprio richiesto per poter operare; se l’impresa scende sotto tale livello, incorre nell’intervento del Regolatore, in Italia l’IVASS - “*Istituto per la Vigilanza sulle Assicurazioni*” - , il quale le imporrà di colmare il disavanzo tramite un apporto di capitale da parte dei soci, entro i tempi stabiliti. È definito inoltre un requisito di capitale minimo: il “*Minimum Capital Requirement*” (MCR), come percentuale del primo, sotto il quale invece l’impresa subisce la revoca diretta della licenza per operare, oltre alla possibile ingerenza nella governance da parte dell’istituto di vigilanza, procedura detta di “*Amministrazione Straordinaria*”.

Per quanto riguarda la sua determinazione, la Direttiva si esprime così:

---

<sup>6</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (Solvibilità II) (25 novembre 2009), art. 75, punto 1, lettere a) e b).

*“Il requisito patrimoniale di solvibilità corrisponde al valore a rischio dei fondi propri di base dell’impresa di assicurazione o di riassicurazione soggetto ad un livello di confidenza del 99,5 % su un periodo di un anno.”<sup>7</sup>*

Viene fatto cioè specifico riferimento all’uso del “Valore a Rischio”, o “Value at Risk” (VaR), ad un livello di confidenza del 99,5%, per la determinazione dell’SCR. Il VaR è una tra le misure di rischio più note ed utilizzate, poiché semplice da calcolare e applicabile ad una qualsiasi distribuzione, ed è così definita: data  $F$  una Funzione di Ripartizione qualsiasi, data  $\alpha \in [0, 1]$ , si definisce il Valore a Rischio a livello  $\alpha$  di  $F$  come:

$$VaR_\alpha[F] = \inf \{x \in \mathbb{R} : F(x) > \alpha\} .$$

Se  $F$  è una funzione invertibile su  $\mathbb{R}$  - come una Normale - o restringendo il dominio all’insieme su cui risulta invertibile - come una Esponenziale in  $[0, +\infty[$  -, per  $\alpha \in ]0, 1[$ , si può dimostrare che:

$$VaR_\alpha[F] = F^{-1}(\alpha) .$$

Ovvero risulta equivalente al quantile di ordine  $\alpha$  della distribuzione  $F$ <sup>8</sup>. L’SCR è pertanto ottenibile come il percentile di ordine 99.5% della distribuzione della variazione dei Fondi Propri di Base in un anno fiscale, presa con segno negativo. L’idea di fondo, rispecchiata dall’utilizzo di questa misura di rischio, è quella di imporre ad una impresa di allocare capitale proprio atto a garantire a questa di essere solvente su un arco di almeno un anno, con un certo livello di probabilità. Con “solvente” si intende, dal punto di vista quantitativo e probabilistico, l’evento: “Il Capitale Proprio risulta positivo alla fine dell’anno” o ugualmente: “Gli Attivi risultano maggiori delle Passività alla fine dell’anno”. In termini matematici il ragionamento, semplificato, può essere costruito nel modo seguente: dato  $N(t)$  il valore ad epoca  $t$  del Capitale Proprio, con FdR  $F_{N(t)}$ , per semplicità supposta invertibile su tutto  $\mathbb{R}$ , fissato  $\alpha$  un certo livello di probabilità vicino ad uno - ad esempio 99.5% - , si vuole calcolare quel capitale  $C \in \mathbb{R}$ , il quale, se allocato al Capitale Proprio dell’anno successivo, garantisca che la probabilità di solvenza sia pari ad  $\alpha$ , ovvero:

$$P(N(t+1) + C \geq 0) = \alpha \text{ sse } P(N(t) + \Delta N(t) + C \geq 0) = \alpha \text{ sse}$$

$$P(-\Delta N(t) \leq N(t) + C) = \alpha \text{ sse } F_{-\Delta N(t)}(N(t) + C) = \alpha \text{ sse}$$

$$SCR = N(t) + C = F_{-\Delta N(t)}^{-1}(\alpha) = VaR_\alpha[-\Delta N(t)],$$

---

<sup>7</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (Solvibilità II) (25 novembre 2009), art. 101, punto 3.

<sup>8</sup>La definizione è in realtà impropria, perché il VaR, così come le misure di rischio, è definito su uno spazio di probabilità per un dato numero aleatorio. Tuttavia, una proprietà fondamentale che si può dimostrare è l’“invarianza per distribuzione”, ovvero dipende solamente dalla FdR del numero aleatorio e non dal numero aleatorio in sé.

Dove  $\Delta N(t) = N(t+1) - N(t)$ <sup>9</sup>. Il livello del Requisito di Solvibilità è di conseguenza ottenuto così come era stato descritto sopra; inoltre, è rappresentabile come somma tra  $N(t)$  e  $C$ : quando  $N(t) > SCR$  - impresa che produce profitti - , allora  $C < 0$  e ciò significa che la Compagnia dispone di un surplus di capitale, pari proprio a  $-C$ .

In questo ragionamento si può vedere perfettamente espresso il tema sopra discusso della valutazione dell'aleatorietà intrinseca nella dinamica patrimoniale, e soprattutto nella valutazione dei rischi che ne determinano questa natura: si vuole garantire agli assicurati, o in generale a tutti gli *stakeholder*, che i benefici e i risarcimenti promessi saranno pagati anche in condizioni estreme e/o inaspettate. Sebbene il VaR sia citato espressamente, il Legislatore incoraggia l'utilizzo di misure di rischio più prudenti al medesimo livello di confidenza, come il "*Tail Value at Risk*" (TVaR), che tiene conto del comportamento della coda della distribuzione dopo il VaR ed è sempre - quando definito - maggiore o uguale a questo, risultando quindi in una allocazione di capitale maggiore.

Passando alle Riserve Tecniche, queste sono la voce più rilevante delle Passività di un'assicurazione e rappresentano una stima degli impegni presi dalla compagnia nei confronti dei propri assicurati. La direttiva definisce la sua valutazione ritagliandola il più possibile sul concetto, sopra definito, di *Market Consistent*:

*“Il valore delle riserve tecniche corrisponde all'importo attuale che le imprese di assicurazione e di riassicurazione dovrebbero pagare se dovessero trasferire immediatamente le loro obbligazioni di assicurazione e di riassicurazione ad un'altra impresa di assicurazione o di riassicurazione.”*<sup>10</sup>

Questo valore, come si può intuire, non è osservato su nessun mercato reale ma nasce piuttosto da uno scambio idealizzato tra due parti, in cui una si assume i debiti dell'altra sotto pagamento di un certo importo, il prezzo dell'operazione. La definizione ne individua il carattere generale, il Legislatore si è poi preoccupato di darne la dimensione particolare: esse sono calcolate come somma di due quantità, ovvero il valore attuale atteso delle passività e attività assicurative e riassicurative - prima quantità -, detta "*Best Estimate of Liabilities*" (BEL), poiché si deve basare sulla miglior stima (in termini predittivi) dei valori futuri, soprattutto senza termini prudenziali, più un margine per il rischio, il "*Risk Margin*" - seconda quantità, in breve RM -. Questo secondo addendo si riallaccia all'idea base della definizione generale: se la seconda impresa decide di assumersi il debito, diviene soggetta a tutti i requisiti patrimoniali previsti dalla direttiva, in particolare all'obbligo di immobilizzare capitale proprio atto a soddisfare il requisito di solvibilità presente e futuro per portare a estinzione la passività, immaginandola riferita ad un portafoglio di assicurati chiuso a nuovi ingressi; quindi,

---

<sup>9</sup>Come detto, il ragionamento è semplificato. È possibile porsi in una condizione più generale, considerando una FdR qualsiasi e tenendo conto dei rendimenti sul capitale allocato.

<sup>10</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (Solvibilità II) (25 novembre 2009), art. 76, punto 2.

il prezzo della transazione deve considerare il costo dell’allocazione dell’SCR. Il valore attuale di tale costo, tenendo conto della remunerazione per il rischio, è il suddetto RM. La procedura permette di ottenere una valutazione delle Passività Assicurative pur non avendo un mercato di riferimento per queste e mantenendosi fedele al concetto di *Market Consisten*. Questo tipo di metodi è detto “*Market to Model*” poiché prevedono di ricavare il valore di mercato di un asset a partire da un modello matematico, non potendo effettuare la valutazione “*Market to Market*”, ovvero, confrontando l’asset con un mercato reale.

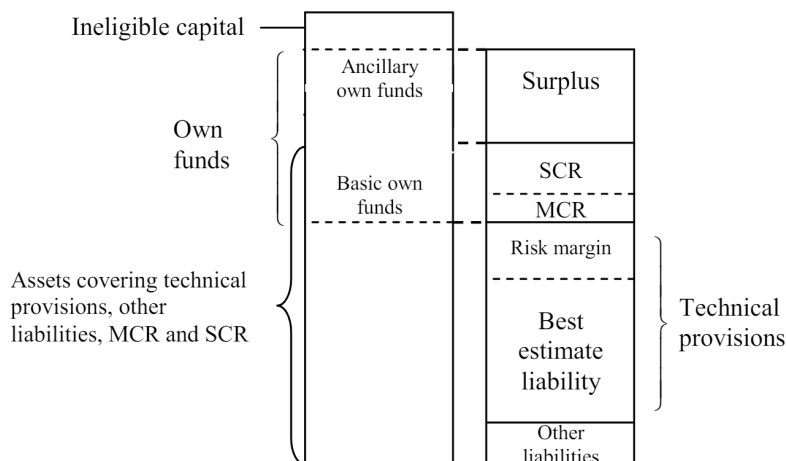


Figura 2: Lo stato patrimoniale nella divisione prevista in *Solvency II*. Fonte: *SOLVENCY II – LIFE INSURANCE* (2016).

### 1.1.2 Il secondo pilastro:

*Solvency*, sul tema della *governance*, scrive:

“*Gli Stati membri richiedono a tutte le imprese di assicurazione e di riassicurazione di dotarsi di un sistema efficace di governance, che consenta una gestione sana e prudente dell’attività.*”<sup>11</sup>

Tale sistema viene a realizzarsi tramite la definizione delle quattro funzioni fondamentali, descritte in questo secondo pilastro, di cui la compagnia si deve dotare, da affiancare al consiglio di amministrazione. Rappresentano dei requisiti qualitativi o formali, che predispongono degli organi di controllo sull’attività assicurativa.

La prima è la funzione di “*Revisione interna*” o di “*Internal Audit*”, che si preoccupa di supervisionare l’intero processo aziendale, fare verifiche sulla correttezza, efficacia ed

<sup>11</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (*Solvibilità II*) (25 novembre 2009), art. 41.

efficienza delle singole procedure e individuare e valutare possibili rischi che ne minacciano la stabilità. Esse deve avere un carattere di assoluta indipendente dal processo gestionale ed esecutivo dell'impresa, per poter essere obbiettiva nelle sue valutazioni.

Segue la funzione di “*Gestione del rischio*” o di “*Risk Management*”. Essa mette in atto le tecniche matematico-statistiche di ERM, in particolar modo si occupa della definizione dell'ORSA: “*Own Risk and Solvency Assessment*”, lo strumento che permette di valutare i rischi interni ed esterni che minacciano l'impresa, e la complessiva solvibilità.

Si ha poi la funzione di “*Controllo interno*” o di “*Compliance*”, che valuta il rischio di non conformità alla normativa, quindi un controllo di natura giuridico sulle procedure, compresa la valutazione degli impatti economici derivanti da rischi legali. Come per la funzione di *Internal Audit*, anche questa deve avere un carattere di assoluta indipendenza.

In fine, si ha la funzione “*Attuariale*”, che va a sostituire la precedente figura dell'attuario incaricato unico. Il ruolo è suddiviso in un attuariato di primo livello, interno all'impresa, che si occupa del calcolo diretto di tutte le componenti di natura statistico-attuariale, come le Riserve Tecniche, in un attuariato di secondo livello, il quale può essere esternalizzato e svolge il ruolo di validazione indipendente dei calcoli svolti dal primo, ed in fine un attuariato di terzo livello, esterno all'impresa, che ha il compito di sorveglianza, consulenza e ulteriore validazione della gestione attuariale interna e dell'operato.

### **1.1.3 Il terzo pilastro:**

Il Regolatore ha ben deciso di introdurre, a cerniera, tutta una serie di obblighi informativi e di trasparenza. Se il secondo pilastro si preoccupa di assegnare dei ruoli interni all'impresa che svolgano il ruolo di supervisori indipendenti, il terzo introduce una forma di controllo esterno, dato non solo da un istituto nazionale preposto - in Italia, come già accennato, vi è l'IVASS, a cui fa capo l'EIOPA, a livello europeo - , ma anche dal mercato stesso. Infatti, la compagnia è obbligata a fornire periodicamente tutta una serie di informazioni interne sotto forma di documentazione scritta, esaustiva, trasparente e chiara, ai fini di far conoscere a tutti gli *stakeholders* il reale stato di salute dell'impresa.

## **1.2 Il Requisito Patrimoniale di Solvibilità**

Nel paragrafo precedente si è visto come il Requisito di Solvibilità possa essere ottenuto come il *Valore a Rischio* della variazione del Capitale Proprio, e come il calcolo del VaR dipenda dalla Funzione di Ripartizione di quest'ultima. Tuttavia la valutazione



di tale FdR è molto complessa da operare in pratica, ciò perché la stocasticità del Capitale Proprio dipende significativamente da un numero rilevante di variabili, le quali, agendo congiuntamente, rendono di fatto molto ostica la costruzione di un modello e la sua stima. Ad ognuna di queste variabile è solitamente associato un particolare rischio, rappresentabile tramite un numero aleatorio o un processo stocastico, a cui è esposta l'assicurazione. Ad esempio, la mortalità può essere vista come un fattore di rischio rilevante poiché da essa dipende in maniera rilevante il risultato della gestione di una compagnia assicurativa che operi in ambito Vita.

La normativa allora, per alleggerire l'onere della stima dell'SCR, ha stabilito quattro approcci di calcolo: l'approccio tramite la “*Formula Standard*” con parametri di mercato, l'approccio tramite formula standard ma con parametri interni all'impresa, gli “*Undertaking Specific Parameters*” (USP), l'approccio tramite costruzione di un “*Modello Interno*”, e in fine un approccio intermedio, che prevede la costruzione di un “*Modello Interno Parziale*”, ovvero che predispone di applicare la formula standard per la determinazione di una quota del requisito e un modello interno per la determinazione della restante quota.

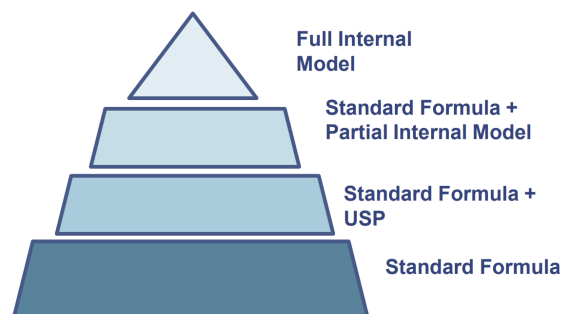


Figura 3: Gli approcci per la determinazione dell'SCR, in ordine di complessità di implementazione: in basso il più complesso, in alto il più semplice. Fonte: Roberti (2017).

La formula standard si configura come un metodo codificato e semplificato, che viene solitamente adottato dalle compagnie quando i costi per l'implementazione del modello interno sono ritenuti o non sostenibili o sproporzionati rispetto al vantaggio che si trarrebbe dall'adozione di un modello interno. Essa infatti risulta molto più semplice da adottare rispetto al modello interno, ma di contro è molto generica e quindi non adattabile al reale profilo di rischio dell'impresa. La procedura che segue la formula standard prevede di suddividere il calcolo dell'SCR in singoli requisiti riferiti a singoli fattori di rischio. Ad ognuno di questi è assegnato un metodo per determinare il cosiddetto “*scenario stressato*”, ovvero uno scenario nel quale il rischio si è manifestato con una certa severità fissata, tale da causare un impatto più facilmente stimabile sulle Attività e Passività della compagnia. A questo punto il singolo requisito di capitale viene ottenuto sottraendo dal Capitale Proprio nello scenario base, tipicamente il valore atteso, il Capitale Proprio nello scenario stressato; la procedura è talvolta denomina-

ta di “*Stress Testing*”. Ogni valore ottenuto è poi aggregato con gli altri tramite una formula, che tiene conto delle possibili correlazioni tra i rischi, determinando così il “*Basic Solvency Capital Requirement*” (BSCR). A questo viene sommato il requisito per il “*Rischio Operativo*”, calcolato separatamente rispetto agli altri poiché in parte comprendente la valutazione del rischio inerente al processo stesso di calcolo dell’SCR, ed il risultato è poi aggiustato per l’effetto di assorbimento delle Riserve Tecniche e delle imposte future negli scenari stressati - c.d. “*loss absorbing capacity of technical provisions and deferred taxes*” - , ottenendo così l’SCR completo<sup>12 13</sup>.

Di fatto, la formula standard sorvola il problema della stima della distribuzione del Capitale Proprio, poiché assegna direttamente una formula per calcolare in modo approssimativo un VaR riferito ad un solo rischio. La somma dei requisiti, pesati per le correlazioni, risulterà in una approssimazione del VaR richiesto nella definizione data dall’articolo 101 della Direttiva<sup>14</sup>.

La procedura appena descritta è usualmente rappresentata tramite una struttura gerarchica detta “Albero dei Rischi di Solvency”, nella quale ogni nodo è detto modulo e ogni foglia è detta sotto-modulo. Vi sono quattro moduli a cui sono associati quattro macro-rischi: “*Rischi di sottoscrizione*”, “*Rischi di Mercato*”, “*Rischi di Default*” e “*Rischi per Attività Immateriali*”. I primi sono i rischi legati alla sottoscrizione di contratti assicurativi con gli assicurati, e vengo distinti in tre moduli: “Vita”, “Non Vita”, e “Salute”, per i rispettivi tre rami dell’assicurazione. Questi tre sono costituiti da sotto-moduli, che rappresentano i rischi base su cui poter calcolare gli scenari stressati. Si possono trovare, senza elencarli tutti, il “*Rischio di Premio*” e il “*Rischio di Riservazione*”, fondamentali per il ramo Danni, e il “*Rischio di Mortalità*” e il “*Rischio di Longevità*”, tipici invece del ramo Vita e Salute. In particolare poi, sempre in questo modulo, si trova il “*Rischio di Estinzione Anticipata*”, rischio legato alla problematica centrale che fa da sfondo alla presente Tesi<sup>15</sup>. Tornando ai moduli restanti, il *Rischio di Mercato* fa riferimento a rischi propri dei mercati, come il “*Rischio di Cambio*” o il “*Rischio di Tasso*”, il *Rischio di Default* invece riguarda la possibile inadempienza, completa o parziale, dei debitori, che grava sulle obbligazioni e sui crediti dell’impresa, ed in fine i *Rischi per Attività Immateriali* sono i rischi legati alle attività immateriali del patrimonio dell’impresa.

---

<sup>12</sup>«*rischio operativo*», il rischio di perdite derivanti dall’inadeguatezza o dalla disfunzione di procedure interne, risorse umane o sistemi, oppure da eventi esogeni”, cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (Solvibilità II) (25 novembre 2009), art. 13, punto 33.

<sup>13</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (Solvibilità II) (25 novembre 2009), art.108.

<sup>14</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (Solvibilità II) (25 novembre 2009), art.101.

<sup>15</sup>Il “*Rischio di Estinzione Anticipata*” non è proprio solo del modulo Vita, ma anche dei moduli Danni e Salute. Il fenomeno è lo stesso ma ha impatti differenti. La presente Tesi, tuttavia, si concentrerà solo sul primo.

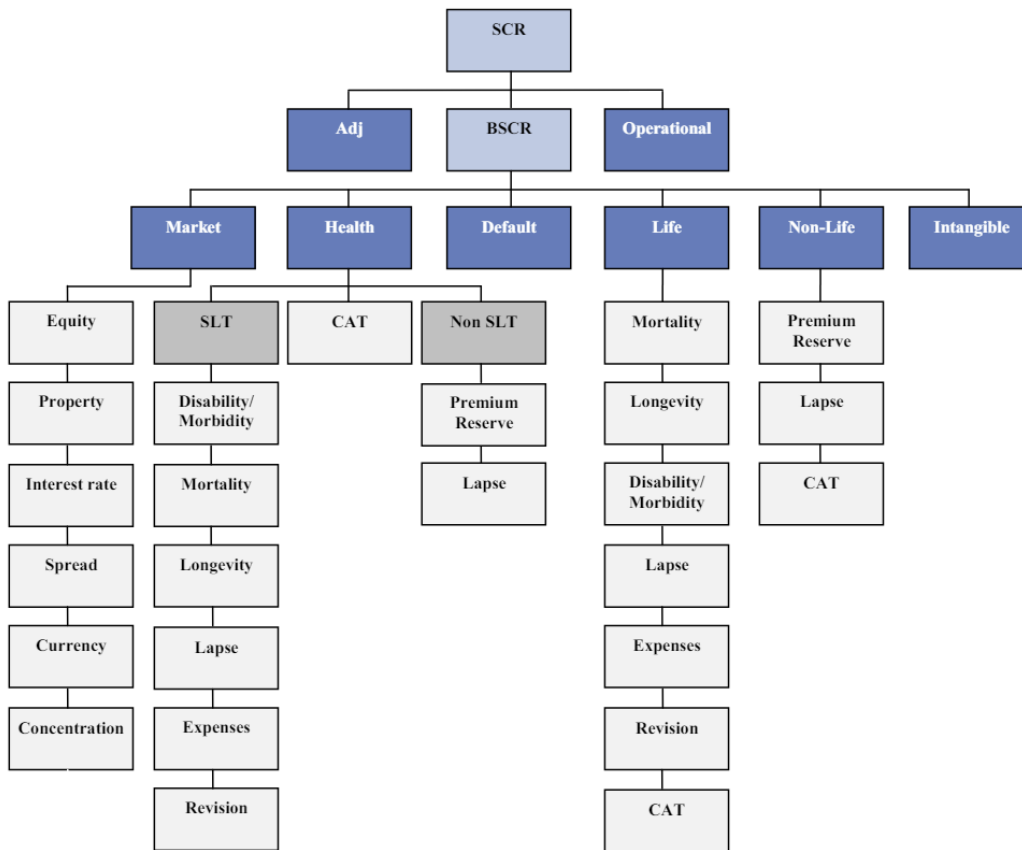


Figura 4: L'Albero dei rischi di Solvency. Fonte: *SOLVENCY II La Nuova Regolamentazione Prudenziale Del Settore Assicurativo: Una Guida Semplificata* (2016).

Come osservato sopra, il metodo non è specifico per una data assicurazione, le quali posso manifestare differenti profili di rischio, ma si presenta il più generale possibile. Ciò fa sì che si possano verificare situazioni nelle quali il requisito calcolato per alcuni rischi da una data compagnia possa risultare inadeguato, o per eccesso o per difetto<sup>16</sup>. Per questo la Direttiva propone altre strade per migliorare l'adattamento, prima fra queste l'utilizzo di parametri interni specifici, gli USP, come indicato sopra, che possono essere utilizzati, se approvati dall'istituto di vigilanza preposto, a sostituzione dei parametri indicati dalla direttiva, quindi più calibrati sugli effettivi rischi a cui è esposto l'assicuratore. Sono ad esempio utilizzati nelle formule per ottenere lo scenario stressato, poiché queste possono dipendere da misure di volume, oppure nella formula di aggregazione dei singoli requisiti, che invece dipende da una valutazione delle correlazioni tra i rischi.

Il massimo grado di libertà ammesso è però raggiunto tramite l'implementazione di un proprio approccio per la determinazione del requisito patrimoniale, sempre rispettando la sua definizione; si parla appunto del modello interno. Esso dev'essere approvato dall'autorità vigilante, la quale conduce dei severi test di qualità e sicurezza del modello, ripetuti anche successivamente all'approvazione. Per quanto più libero, la struttura del modello interno deve comunque essere impostata in modo tale da stimare la distribuzione del Capitale Proprio, tenendo conto di tutti i rischi che la Direttiva indica - sostanzialmente gli stessi definiti nella formula standard all'interno dell'Albero dei Rischi - , nello specifico deve poter rispondere a queste domande per ogni rischio elencato: “*Qual è la distribuzione dell'impatto annuo stimato?*”, “*Qual è lo scenario estremo annuo considerato?*” e “*Qual è la frequenza annua di questo scenario estremo?*”. In fine deve poter integrare assieme tutte le valutazioni tenendo conto dei possibili effetti di interazione e correlazione.

Un modello interno si compone di solito di una “Funzione di Aggregazione”, dipendente dall'insieme dei rischi previsti per Direttiva e dalle loro distribuzioni - componente stocastica - , più un insieme di parametri o “Variabili Decisionali” - componente discrezionale - , che dipendono dalla scelta dall'assicuratore. Per determinare il requisito di capitale si può poi procedere in più modi, ad esempio seguendo il metodo della formula standard, ovvero facendo degli stress test per i singoli rischi, determinando quindi uno scenario stressato, secondo un certo criterio, e stimando l'impatto di questo sul Capitale Proprio, tramite la differenza descritta sopra. Bisogna però ripetere che qualsiasi procedura seguita dev'essere in grado di rispondere alle domande sopracitate e soprattutto di poter asserire in fine che il requisito rispetta la definizione data.

Per anticipare il tema della Tesi, un esempio di rischio per il quale si è fatto uso della formula standard nel modello interno è stato, per molte imprese e per molto tempo, il rischio di estinzione anticipata ed in particolare dell'estinzione anticipata di massa, poiché nei periodi precedenti l'innalzamento dei tassi, il rischio determinava un livello di requisito molto basso, a volte trascurabile.

---

<sup>16</sup>In effetti è possibile che l'istituto di vigilanza preposto imponga all'impresa di dotarsi di un modello interno se ritenesse il suo profilo di rischio totalmente inadeguato per poter utilizzare la formula standard.

La creazione e l'implementazione del modello interno sono evidentemente molto complessi. Di fatto non sono altro che più modelli messi assieme, i quali richiedono ognuno giustificazioni di ipotesi non scontate, come ad esempio la scelta della distribuzione e la sua stima, e, problema ancora peggiore, richiedono di definire le interazioni tra le variabili. La stessa complessità porta ad avere molto spesso l'impossibilità ad ottenere risultati in forma chiusa. L'approccio finale è sempre di tipo simulativo, utilizzando metodi Monte Carlo per le stime, e il tempo impiegato per svolgere la simulazione su una macchina è spesso molto lungo, troppo elevato per poter ripetere la procedura più volte di seguito. Anche per questo le imprese si direzionano, almeno inizialmente, verso un modello parziale, per andare a semplificare, utilizzando la formula standard, lì ove possibile e conviene. La compagnia deve appunto fare una valutazione di quanto possa convenire, in termini di costo opportunità, l'implementazione di un modello per un dato rischio, che potrebbe adattarsi meglio al profilo di rischio e possibilmente ridurre il requisito, pagando però la sua costruzione e il mantenimento, oppure utilizzare direttamente ciò che già è fornito da *Solvency*, trovandosi di fronte ad un *trade-off* tra complessità e adattabilità.

### 1.3 Le Riserve Tecniche in ambito Vita

Le Riserve Tecniche, come già visto, si compongono della *Best Estimate of Liabilities* e del *Risk Margin*. Il loro valore costituisce non solo una passività patrimoniale, ma anche il livello di accantonamento dei premi netti incassati dai contratti e dei relativi rendimenti necessari a far fronte agli impegni attesi contrattuali, rappresentanti il debito verso gli assicurati<sup>17</sup>. Tale importo viene investito in una serie di attività più meno liquide e a bassa rischiosità, le quali andranno a costituire l'insieme delle attività poste a copertura delle riserve. Queste attività generano a loro volta dei rendimenti: si parla infatti di "*Rendimenti sulla Riserva*"<sup>18</sup>. In ambito Vita, dove la componente finanziaria è elevata, oltre alla garanzia base, si può assegnare ad un contraente una parte di questi rendimenti, proporzionatamente alle dimensioni e alla tipologia del contratto sottoscritto, in modo tale da incrementare l'importo dei benefici.

Questa forma di partecipazione agli utili prende in Europa tipicamente tre forme:

- distribuzione di profitti maturati;
- distribuzione di profitti futuri;
- distribuzione di proventi finanziari maturati.

---

<sup>17</sup>Le Riserve non sono l'unica voce delle passività nello stato patrimoniale di una assicurazione, ma ne sono la più importante.

<sup>18</sup>Non è propriamente la riserva a generare montante, dato che si tratta di una voce del passivo, ma le sue attività a copertura.

Una fondamentale differenza tra questi sta nel fatto che, nei primi due casi, il profitto è calcolato al netto della tassazione, poiché riferito al risultato economico, mentre l'ultimo è al lordo. Suddetto caso è quello classico italiano e si parla più propriamente di “*Prodotti Rivalutabili*”, e la gestione dei rendimenti interni assegnabili a detti prodotti è indicata col termine “*Gestione Separata*”, o in inglese “*Segregated Fund*”.

Emerge da questo meccanismo di distribuzione degli utili una dinamica sottostante che interconnette attività e passività, investimenti e finanziamenti: i rendimenti sugli attivi vengono retrocessi agli assicurati, diventando parte dei passivi e incrementando la riserva; gli incrementi sulla riserva impongono l'allocazione di nuovi attivi. Ne consegue che non si possono considerare separatamente le due sezioni del bilancio nelle valutazioni tecnico-attuariali.

È importante sottolineare che non tutte le attività risultano adeguate alla copertura ai fini Solvency. La direttiva, infatti, delinea un regolamento fitto per definire le cosiddette “*Attività Ammissibili*”, oltre ad imporre negli obblighi discrezionali di redigere un documento contenente le informazioni sugli investimenti e sulle azioni presenti e future che l'impresa intende mettere in atto, riguardanti la dinamica patrimoniale sopraccitata. Inoltre, viene richiamato esplicitamente il “*Principio della Persona Prudente*” - in inglese “*Prudent Person Principle*” - come *modus operandi* della gestione degli investimenti: vi è libertà nella scelta delle attività in cui investire - limitatamente ai criteri delle Attività Ammissibili - , purché sia affiancato un robusto comparto di *Risk Management*, non a fini speculativi, ma per garantire i rendimenti previsti avendo rischi minimi<sup>19</sup>. Tra le azioni consigliate vi sono quelle soprattutto di riduzione della volatilità, come il “*matching*” tra entrate ed uscite ed investimenti a lungo termine. Le Attività Ammissibili sono costituite in gran parte da titoli obbligazionari, solitamente titoli di stato o titoli emessi da imprese a basso rischio di credito o di default, e una minima parte di *Equity*, sempre però valutati a basso rischio. Oltre a questi, sono ammessi derivati finanziari, che però trovano una normativa ancora più stringente, e devono essere adoperati solo come strumenti di *hedging*. Parimenti sono ammessi strumenti derivati dalla cartolarizzazione, ad esempio di mutui, anche qui soggetti a stringenti controlli e normative - soprattutto in ricordo della crisi finanziaria del 2008 - .

Tornando alle Riserve Vita nel settore italiano, a causa di questa dinamica di partecipazioni, la determinazione delle BEL risulta essere altamente complessa, poiché diventa necessaria la valutazione delle garanzie, implicite ed esplicite, di un contratto e la valutazione dei rendimenti futuri di cui il contraente gode in aggiunta alla garanzia, per effetto della partecipazione. Inoltre, se i contratti prevedono il diritto di riscatto, si dovrà tener conto del valore di riscatto e del comportamento del contraente stesso<sup>20</sup>. Andando più nel dettaglio: per ogni contratto Vita del portafoglio viene stimata la proiezione dei flussi futuri riconducibili al perimetro del contratto, nei diversi possibili

---

<sup>19</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (*Solvibilità II*) (25 novembre 2009), art. 132.

<sup>20</sup>La questione sarà meglio discussa più avanti nel capitolo riguardante i riscatti.

scenari, i quali comprendono in particolare: l'evento morte o sopravvivenza, recesso dal contratto ed evoluzione dei tassi e rendimenti, e ogni combinazione viene pesata per la probabilità corrispondente. Si potranno avere per esempio scenari in cui i rendimenti superano le garanzie e quindi, nel caso in cui si verificano, si dovrà disporre di capitale sufficiente per pagare i benefici aumentati per effetto della partecipazione, ed altri scenari in cui i rendimenti scendono sotto le garanzie, di fatto generando perdite o utili, poiché l'assicuratore dovrà andare a colmare il disavanzo, smobilitando asset, con la possibilità di realizzare minusvalenze o plusvalenze di bilancio<sup>21</sup>.

Ambo gli scenari, di rendimenti maggiori o minori delle garanzie, risultano tanto rilevanti agli occhi del legislatore che ha deciso di imporre che alla voce di bilancio delle BEL venissero scorporare due voci che potessero mettere a risalto l'impatto della semplice garanzia e l'impatto del sovra rendimento possibile<sup>22</sup>. Le due voci sono le "Riserve per le Garanzie Minime" o "Minimum Guaranteed Provisions"(MGP) e le "Riserva per Futuri Benefici Discrezionali" o "Future Discretionary Benefits"(FDB). L'utilità della scomposizione della BEL risiede anche nel fatto che la capacità di assorbimento delle riserve tecniche, ovvero l'effetto di riduzione del requisito patrimoniale tramite il decremento dei benefici discrezionali, è dipendente dal valore del FDB<sup>23</sup>.

La valutazione di queste quantità e in particolare del valore temporale della garanzia, il cosiddetto "Time Value of Options and Guarantees", risulta essere una operazione non immediata, non soltanto a causa dei problemi legati alla stima delle distribuzioni, ma anche al calcolo del prezzo della garanzia sottostante, la quale deve rispettare vincoli finanziari, trattandosi a tutti gli effetti di un prodotto finanziario, in particolare quasi sempre una opzione "camuffata"; in questo senso ci si riallaccia al tema della valutazione *market consistent* di *Assets* e *Liabilities*.

Altro problema sopraccitato è legato alle azioni - legittime - compiute dal contraente, anche queste dipendenti dalle garanzie contrattuali. Tra le varie azioni, la prima fra tutte è la possibilità di recedere dal contratto ricevendo una somma di denaro detta "Valore di Riscatto", la "Surrender Option", quando la polizza preveda in ogni caso il pagamento di un beneficio - ad esempio la mista - . Ognuna di queste azioni ha un costo ed una rilevanza nella determinazione degli scenari. Perciò, sempre richiamandosi al tema della Tesi, è necessario poter valutare e stimare il comportamento futuro dell'assicurato; l'argomento sarà però meglio sviscerato nel capitolo successivo.

---

<sup>21</sup>Per quanto sembri una posizione positiva, quella dei rendimenti più alti delle garanzie, non è sempre detto si disponga effettivamente di tale somma da attribuire. Questo può accadere in prodotti come *Unit Linked* o simili, nei quali può sorgere il problema della replicabilità dei flussi quando l'assicuratore non posseda in tutto o in parte le quote previste per contratto del fondo.

<sup>22</sup>"Future cash-flows also need to be split into guaranteed and discretionary benefits because, as stated in Article 108 of the Level 1 text, the loss absorbing capacity of technical provisions is limited by the technical provisions relating to the future discretionary benefits. The risk mitigation effect provided by future discretionary benefits shall be no higher than the sum of technical provisions and deferred taxes relating to those future discretionary benefits." Cfr. *CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: Technical Provisions Article 86 a Actuarial and Statistical Methodologies to Calculate the Best Estimate* (2009), CP 39, Cap. 3, sez. 3.7, punto 3.169.

<sup>23</sup>Ibidem.

Un'ultima problematica che si può menzionare, propria della valutazione della BEL, è la scelta di una appropriata struttura per scadenze dei tassi con cui calcolare i rendimenti futuri attesi e con cui attualizzare gli importi. Nell'ambito della Direttiva la curva è fornita dall'EIOPA e la struttura implicata è da considerarsi *risk-free*. Essa è stimata a partire dagli *SWAP rates* osservati su mercati sufficientemente liquidi, quindi relativamente privi di rischio. Sono in realtà subentrate successivamente delle modifiche alla normativa che permettono l'utilizzo di alcuni particolari tipi di *spread* da applicare al tasso.



## 2 Riscatto e Abbandono

### 2.1 Il diritto di Riscatto e Abbandono e altri diritti nelle polizze vita

La polizza assicurativa, come tutti i contratti, prevede una serie di diritti e doveri da ambo le parti. Nello specifico, ponendosi dal lato del contraente/assicurato, questi ha il dovere di pagare i premi concordati alle date concordate e il diritto a ricevere, in caso si verificasse l'evento assicurato, il risarcimento, in importo e forma previsti dal contratto, o il beneficio concordato, a seconda che si tratti di polizza vita o danni. In aggiunta a tale diritto, definibile "basilare", egli può disporre di ulteriori diritti, concordati a stipula a sua discrezione, oppure conferitigli per legge, ai quali fanno da controparte altri oneri.

Le tipologie su cui ci si soffermerà sono proprie delle assicurazioni sulla vita e prendono la forma di modifiche unilaterali del contratto, ovverosia il contraente, esercitando il diritto, modifica le condizioni contrattuali e conseguentemente le prestazioni future. Si parla ad esempio di diritto di *interruzione* - detto "*Riduzione*" della polizza - o *variazione*, incrementativa, decrementativa o mista, dei premi futuri, che dal punto di vista attuariale equivale ad una polizza a premi unici ricorrenti; o ancora, di diritto al *rinnovo* del contratto a condizioni prestabilite o viceversa diritto alla *chiusura* del contratto senza rinnovo, per contratti che prevedano il rinnovo tacito, detto diritto di "*Disdetta*"; un altro diritto è quello di prelevare capitale dalla polizza, riducendo il beneficio, il cosiddetto "*Prestito su Polizza*", nel caso in cui sia prevista una prestazione dell'assicuratore certa<sup>24</sup>. Quello più rilevante è però il diritto di "*Estinzione Anticipata*" del contratto, ovvero la possibilità per il contraente di chiudere la polizza - a volte detto "*Recesso*" del contratto assicurativo - , e quindi la copertura, prima del termine di essa<sup>25</sup>. Suddetto diritto ricade sotto due forme, a seconda che vi sia o no una attribuzione di una somma di denaro in caso di recesso; si parla infatti più propriamente di "*Riscatto di una Polizza*" quando questa presenti tale attribuzione, e il capitale viene chiamato "*Valore di Riscatto*". Il diritto di riscatto è sempre presente per legge quando sia prevista una prestazione certa dell'assicuratore nei confronti dell'assicurato, come

---

<sup>24</sup>Certa nel senso di sicuramente dovuta, ma non è noto quando verrà pagata, come nel caso di assicurazioni Caso Morte a vita intera o di assicurazioni Miste, come già accennato nel primo capitolo, nella sezione riguardante le Riserve Tecniche. Bisogna aggiungere che non è possibile l'esercizio nei primi tre anni, a volte due, di vita della polizza per garantire la formazione della riserva matematica del contratto e coprire le spese di acquisizione.

<sup>25</sup>Cosa che invece non accade nel caso della semplice interruzione dei premi, la quale prevede unicamente la riduzione del beneficio in caso si verifichi l'evento assicurato.

nel caso del prestito su polizza. Qualora invece il recesso non presenti l'attribuzione del valore di riscatto si parla semplicemente di “*abbandono*” del contratto. Le forme di garanzie elencate posso anche dipendere dal tipo di contratto/prodotto a cui ci si riferisce. Lo si può vedere considerando ad esempio prodotti come le *Unit Linked* e le *Index Linked*, e soprattutto le *Variable Annuity*, una tipologia di contratti assicurativi altamente modificabili a discrezione del contraente, categorizzate appunto tramite il tipo di garanzia offerto.

Un accenno all'argomento e alle definizioni era già stato dato nel capitolo precedente, parlando dell'impatto delle garanzie contrattuali e delle azioni - legittime - degli assicurati sulla valutazione delle Riserve Tecniche. Di fatto queste “azioni” non sono altro che l'esercizio di diritti, mentre il valore di riscatto è un tipo di garanzia, che può prendere la forma di restituzione di tutto o parte dei premi versati, o un montante di questi ad un tasso fisso concordato, o persino comprendere un'ulteriore garanzia di tipo partecipazione agli utili<sup>26</sup>. Tipicamente però è definito come una certa percentuale della Riserva Matematica del contratto o, più precisamente, della riserva “*zillmerata*”, ovvero la riserva al netto delle spese di acquisizione, le quali sono di competenza dell'assicurato, ma sono anticipate dall'assicuratore e ammortizzate poi nel tempo. Costituiscono pertanto dei crediti per l'impresa e nel momento del riscatto, riducendo la riserva e quindi il valore di riscatto, il recedente salda il suo debito.

Quindi, tornando alle riserve, per determinarne il valore è fondamentale poter studiare l'evoluzione temporale del valore garantito e il comportamento degli assicurati, in particolare prevederne le azioni future, un problema ben diverso dal modellizzare fenomeni di tipo fisico. Per di più, tale problematicità ricade anche, come già accennato, sulla valutazione dell'SCR e sul *pricing* della garanzia e quindi della polizza in sé.

In merito al prezzamento però, il problema può essere in parte semplificato se si considera, riprendendo l'osservazione fatta nel capitolo precedente, che una polizza vita presenta un'importante componente finanziaria, la quale richiede per la sua valutazione di rispettare quelle che sono le ipotesi del mercato, in particolare la “*Legge del Prezzo Unico*”; ancora di più se si osserva che un diritto contrattuale legato ad una garanzia non è altro che una opzione, o in generale un derivato. Se si prende per semplicità una polizza che preveda il diritto di riscatto, il possessore di questa potrà esercitare tale diritto negli istanti precedenti il termine naturale del contratto, situazione analoga a quella di una opzione di tipo americana, che permette l'esercizio dell'opzione negli istanti antecedenti la *maturity*<sup>27</sup>. Ecco allora che il valore di riscatto può essere visto come il sottostante del derivato, mentre lo *strike* è interpretabile come la perdita della copertura assicurativa e quindi del beneficio a scadenza e nel caso di decesso<sup>28</sup>. Ne

---

<sup>26</sup>Sono solitamente previste delle penali di uscita, che riducono il valore di riscatto rispetto a queste opzioni elencate, ciò perché, in linea di massima, si vuole spingere sempre l'assicurato a rimanere nel portafoglio; anche se la sua uscita potrebbe portare una riduzione delle Riserve Tecniche, sul lungo periodo vuol dire sempre perdita di clienti.

<sup>27</sup>Come in una nota precedente, a proposito del prelievo su polizza, anche nel caso del riscatto non è possibile l'esercizio nei primi tre anni, in alcuni casi due. Di conseguenza non vi è una perfetta corrispondenza tra l'opzione americana e l'opzione di riscatto.

<sup>28</sup>È sottinteso nel ragionamento che il valore di riscatto può essere reinvestito sul mercato ad un

consegue poi che il prezzo dell'opzione è il valore della garanzia, da cui la necessità di utilizzare le ipotesi di mercato.

Vi è, tra le ipotesi di mercato, un'altra che è necessario citare, perché fondamentale per l'analisi del comportamento dell'assicurato, ovverosia l'ipotesi che il mercato sia popolato da “*agenti razionali e non saziati*”, con la quale si intende che gli operatori del mercato agiscono secondo logiche razionali al fine di ottenere la massimizzazione del profitto. Facendo valere questa ipotesi anche per il caso dei contraenti delle polizze, allora è possibile prevedere se e quando decidano di riscattare, ovvero nell'istante in cui il profitto atteso dal riscatto è maggiore del profitto atteso dalla polizza, così come avviene, almeno approssimativamente, per una opzione americana<sup>29</sup> <sup>30</sup>. Il calcolo è comunque molto complesso, esiste una pletora di articoli scientifici che propongono diversi metodi per il prezzamento dell'“opzione riscatto” e per la determinazione del tempo ottimale di recesso; a questo proposito si può citare l'articolo Bacinello, Biffis, and Millosovich (2010).

Si può tuttavia asserire anche che i contraenti/assicurati non si comportino sempre razionalmente, anche quando ne hanno l'intenzione, poiché potrebbero svolgere valutazioni errate. Per di più, usare il termine “razionale” per riferirsi al caso del recesso può essere fuorviante - sebbene venga usato molto spesso per descrivere il comportamento di alcuni assicurati, da contrapporre a quello di altri assicurati detti “Non Razionali” - perché la razionalità è intesa solamente nella sua definizione data sopra di “*agente razionale*”, quindi per indicare il comportamento di quell'individuo che sceglie di riscattare valutando - qui subentra la razionalità di per sé - la convenienza di reinvestire il valore di riscatto in un diverso strumento e generare un montante superiore al beneficio atteso dal contratto. Un secondo individuo, che si trovasse ad affrontare un problema personale di liquidità, come l'improvvisa ed inaspettata rottura del proprio mezzo, agirebbe razionalmente se andasse a riscattare la polizza per coprire l'esborso qualora ingente, ma dal punto di vista finanziario perderebbe un possibile profitto futuro e perciò la sua scelta non sarebbe comunque considerata “razionale” secondo la definizione data. La teoria che descrive questo tipo di soggetti è detta “*Emergency Fund Hypothesis*” ovvero l'ipotesi del “fondo di emergenza”; il discorso in merito verrà meglio ampliato nell'ultimo paragrafo del presente capitolo. È importante, però, tenere a mente che, mentre per il *pricing* del prodotto si deve considerare unicamente il comportamento in ipotesi di mercato ideale, nel caso della valutazione del rischio di recesso, e quindi dell'SCR o delle RT, si devono - per obbligo esplicito del Regolamento Delegato<sup>31</sup> - invece prendere in esame tutte le possibili variabili che danno origine ai diversi compor-

---

certo tasso, diverso da quello atteso nella polizza. Questo rende il montante a scadenza stocastico, così come avviene in un *Forward Rate Agreement*.

<sup>29</sup>Dato che, come è stato detto, la garanzia di riscatto può essere vista come una opzione, si può ricordare che è proprio questa ipotesi di razionalità e non sazietà degli agenti che permette di esprimere il *payoff* di una opzione come il massimo tra zero e la differenza tra lo *strike* e il valore del sottostante.

<sup>30</sup>Come riferimenti al tema della valutazione dei derivati, e in particolare alla valutazione delle opzioni si rimanda ai libri: Hull and Basu (2016) e Ingersoll (1987).

<sup>31</sup>Cfr. *REGOLAMENTO DELEGATO (UE) 2015/35 DELLA COMMISSIONE* (2014), art. 26.

tamenti. La valutazione del rischio di riscatto condizionata ai soli assicurati razionali può comunque risultare molto utile perché, almeno in via teorica, permette di ottenere un *worst case scenario*, dato che quando l'assicurato riscatta nell'istante di ottimo, l'assicuratore, avendo la posizione opposta nel contratto, ha una perdita, sottolineando però ulteriormente che vale solo in via teorica.

Perché si parla di “Rischio” di estinzione? Dal punto di vista matematico ogni rischio è rappresentabile tramite un numero aleatorio, come può essere visto il tasso di recesso annuo, ma dal punto di vista assicurativo il rischio è legato ad un impatto di tipo monetario - solitamente una perdita - che si manifesta a causa del dato fenomeno aleatorio. È stato detto in precedenza che è necessario, nella valutazione dell'SCR, tener conto del rischio dell'estinzione anticipata, si può andare a vedere ora più nel concreto il motivo, tenendo a mente la definizione del capitale di solvibilità data da *Solvency*<sup>32</sup>. Si prenda come esempio un portafoglio di polizze che prevedano il diritto di riscatto; l'assicuratore accantona parte dei premi per costituire la riserva matematica, valutando in termini attesi i *cash flow* annui, in particolare riferiti ai decessi e ai riscatti. I premi sono quindi investiti in attività prive di rischio, come titoli di stato o obbligazioni ritenute sicure, calibrando le scadenze in funzione del momento in cui sono attese le uscite di cassa, avendo cura di disporre di liquidità per coprire le spese annue attese. È chiaro che ragionando in termini attesi si sbaglia sempre - ma di meno rispetto a qualsiasi altro ragionamento - : vi saranno anni in cui le spese superano il valore atteso e l'assicuratore dovrà vendere obbligazioni sufficienti prima della scadenza, al loro valore di mercato, oppure anni in cui le spese rimangono sotto il valore atteso e si avrà un surplus di liquidità. Si ponga il caso di un incremento inaspettato dei recessi. Vengono smobilitati *asset* prima della *maturity* al loro valore di mercato, il quale potrà essere maggiore o minore del valore di acquisto degli stessi. Se il valore è maggiore il portafoglio realizza delle plusvalenze, di fatto cioè genera un profitto, se invece il valore è minore si realizzano delle minusvalenze, quindi delle perdite.

Si supponga ora, caso più interessante, che vi sia un improvviso aumento dei tassi di interesse, magari per effetto di una decisione della Banca Centrale nell'intento di ridurre un'inflazione troppo elevata. Quando i tassi crescono i prezzi scendono, e le imprese assicurative possono trovarsi da un momento all'altro ad avere un portafoglio obbligazionario che perde di valore, poiché ricolmo di *bond* con tassi di rendimento più bassi del tasso di mercato. Non solo, le polizze hanno dei tassi impliciti che dipendono dai rendimenti di quel portafoglio, quindi rendimenti bassi rispetto al mercato. Questo innesca un processo particolare: gli assicurati più accorti di finanza si rendono conto che il rendimento della polizza è inferiore rispetto alla media del mercato e si trovano quindi incentivati a riscattare le polizze e reinvestire il valore di riscatto in altri titoli più remunerativi. L'assicuratore si ritrova con un aumento improvviso dei riscatti sopra la media, per cui deve svendere titoli, svendere appunto perché questi avranno un valore sicuramente - in questa circostanza - minore di quello d'acquisto. Il risultato è una

---

<sup>32</sup>Il discorso in realtà è valido in generale, non solo nell'ottica della Direttiva, perché il rischio ovviamente non è legato alla normativa vigente, ma all'impresa stessa.

perdita per il portafoglio – almeno in termini di valore di bilancio – che impatterà sui rendimenti e sul Capitale Proprio. Per questo *Solvency* richiede di valutare un livello apposito di accantonamento che possa coprire una possibile situazione di default dovuta a tale situazione.

Si noti che la dinamica qui rappresentata è in effetti quella verificatasi tra il 2022 e il 2023, come verrà raccontato più avanti; in particolare è ciò che è accaduto per il caso Eurovita, riportato nell’Introduzione.

Un commento è necessario: un aumento improvviso del tasso di estinzione non costituisce di per sé una perdita certa, anzi, nel caso in cui il portafoglio presenti plusvalenze non realizzate, l’assicuratore può coprire i valori di riscatto da pagare agli uscenti realizzandole, e conseguentemente generando profitti; lo stesso vale viceversa quando si osservano meno riscatti del previsto: questo di per sé non costituisce un profitto, poiché si potrebbero non realizzare delle plusvalenze che invece ci si aspettava e quindi avere una riduzione dei profitti attesi, oltre a richiedere un aumento delle RT per gli assicurati che ci aspettava uscissero e che invece sono rimasti.

Altro motivo, più intuitivo, per considera il *lapse risk* un rischio effettivo per l’impresa risiede nel fatto che gli assicurati che recedono escono dal portafoglio, ovvero sono clienti persi. I flussi di premi futuri attesi sono di conseguenza ridotti e la numerosità del pool di assicurati cala. Ciò fa sì che si riduca l’effetto di *pooling*: in generale si cerca di costruire gruppi ampi di contratti, il più possibile omogenei e indipendenti, in modo tale che la variabilità relativa attorno alla spesa attesa si riduce, rendendo le previsioni più sicure. Per questo motivo, l’effetto di solidarietà tra gli assicurati ha una maggior efficacia, nel senso che il montepremi incassato riesce meglio a coprire i benefici attesi.

## 2.2 Il rischio di estinzione anticipata nella Direttiva

Considerato quanto detto al paragrafo precedente, risulta evidente il motivo per cui il Legislatore si sia preoccupato di inserire nella direttiva, tra gli elementi da valutare, il suddetto rischio. Si può ora andare a presentare, più nel dettaglio, come *Solvency* definisca il rischio di estinzione e come lo valuti tramite la formula standard.

L’articolo 104 della Direttiva elenca i moduli di rischio che devono comporre il requisito patrimoniale, mentre il successivo, il 105, definisce tali moduli ed in particolare presenta i sotto-moduli che li compongono<sup>33</sup>. Al punto 3 del 105, si parla nello specifico del modulo per i rischi di sottoscrizione per l’assicurazione vita e dei suoi sotto-moduli, tra i quali si trova alla lettera *f* il “*Rischio di Estinzione Anticipata*”, indicato nel testo inglese con il termine “*lapse risk*”; così è come viene definito tale rischio:

---

<sup>33</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (*Solvibilità II*) (25 novembre 2009), art. 104 e 105.

“il rischio di perdita o di variazione sfavorevole del valore delle passività assicurative, derivante da variazioni del livello o della volatilità dei tassi delle estinzioni anticipate, dei recessi, dei rinnovi e dei riscatti delle polizze (rischio di estinzione anticipata);”<sup>34</sup>

Con il termine “*estinzione anticipata*” vengono pertanto raccolti tutti i diritti che sono stati elencati nel paragrafo precedente. Per evitare confusioni nella nomenclatura, si utilizza spesso in italiano il termine “*lapse*” per indicare l’insieme di questi diritti, che è il termine utilizzato non solo dalla Direttiva ma anche in letteratura - non sempre - appunto nella stessa accezione<sup>35</sup>. Viene fatto ciò per ricordare che il requisito si riferisce ad ognuno di essi.

Ancora più importante, uscendo dalla Direttiva quadro e spostandosi al Regolamento Delegato 2015/35/UE, si può trovare la suddivisione del rischio in tre sotto-moduli<sup>36</sup>: rischio di “*Aumento Permanente*” del tasso di estinzione, “*Lapse-up risk*”, rischio di “*Calo Permanente*” del tasso di estinzione, “*Lapse-down risk*”, e rischio di “*estinzione di Massa*”, “*Mass Lapse risk*”<sup>37</sup>. Con i primi due la normativa indica il rischio di sottostima, *lapse-up*, o di sovrastima, *lapse-down*, in termini *best estimate* del tasso di estinzione anticipata; essi sono in particolare connessi ai problemi presentati nel primo esempio del paragrafo precedente. L’ultimo è invece quello propriamente indagato dalla presente Tesi e si riferisce al rischio di un improvviso ed inaspettato aumento estremo del *lapse rate* osservato su un orizzonte temporale di un anno; un evento a tutti gli effetti catastrofico, sebbene la specifica di “*rischio catastrofico*” in ambito vita per *Solvency* si riferisca solo al tasso di decesso. Il secondo esempio del paragrafo precedente si rifà a questo tipo di evento.

Sempre il Regolamento Delegato fornisce la procedura di calcolo e i valori da utilizzare nella formula standard per la determinazione dell’SCR di tale modulo<sup>38</sup>. Il requisito per il rischio di estinzione viene ottenuto come il massimo tra i tre requisiti riferiti ai tre sotto-moduli: il primo è la valutazione dell’impatto sul Capitale Proprio di un aumento permanente del 50% del *lapse rate* atteso, il secondo invece è la valutazione dell’impatto per un calo permanente sempre del *lapse rate* atteso del 50%, tale da produrre una differenza massima di 20 punti percentuali, infine il terzo è distinto tra polizze riferite a fondi pensionistici collettivi, polizze individuali e particolari contratti di riassicurazione che riguardano contratti assicurativi futuri che rientrano nel calcolo delle RT: per il primo gruppo di polizze si valuta l’impatto di una perdita annua del 70% dei contratti, per il secondo gruppo del 40%, mentre per il terzo un calo del 40%

---

<sup>34</sup>Cfr. DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riassicurazione (*Solvibilità II*) (25 novembre 2009), art. 105, punto 3, lettera f.

<sup>35</sup>Per una definizione del termine *lapse* si può fare riferimento ad esempio a: Shaughnessy and Tewksbury (2019), pag. 12.

<sup>36</sup>REGOLAMENTO DELEGATO (UE) 2015/35 DELLA COMMISSIONE (2014).

<sup>37</sup>Cfr. REGOLAMENTO DELEGATO (UE) 2015/35 DELLA COMMISSIONE (2014), art. 142, punto 1.

<sup>38</sup>Cfr. REGOLAMENTO DELEGATO (UE) 2015/35 DELLA COMMISSIONE (2014), art. 142.



dei contratti futuri<sup>39</sup>. Il requisito minimo imposto per normativa è di zero, cioè vengono esclusi i casi in cui il requisito risulta negativo, i quali possono verificarsi quando il valore del Capitale Proprio, valutato negli scenari stressati, risulta maggiore del suo valore nello scenario base, situazione possibile, come visto in precedenza.

Per rendere più chiaro il calcolo, lo si esprime nelle seguenti formule:

$$SCR_{lapse} = \max\{SCR_{up} ; SCR_{down} ; SCR_{mass} ; 0\} ,$$

dove:

$$SCR_{up} = NAV|\{\Lambda = \lambda_{BE}\} - NAV|\{\Lambda = \lambda_{up}\} ,$$

$$SCR_{down} = NAV|\{\Lambda = \lambda_{BE}\} - NAV|\{\Lambda = \lambda_{down}\} ,$$

$$SCR_{mass} = NAV|\{\Lambda = \lambda_{BE}\} - NAV|\{\Lambda = \lambda_{mass}\} ,$$

e dove:

$$\lambda_{up} = \min\{1, 5 \lambda_{BE} ; 1\} ,$$

$$\lambda_{down} = \max\{0, 5 \lambda_{BE} ; \lambda_{BE} - 0, 2\} ,$$

$$\lambda_{mass} = \begin{cases} 0, 7 & \text{per le polizze del primo gruppo} \\ 0, 4 & \text{per le polizze del secondo e terzo gruppo} \end{cases} .$$

Con  $SCR_{...}$  si intende il requisito di capitale specifico per il modulo *lapse* o per i tre sotto-moduli. Questi ultimi tre sono calcolati come differenza tra il Capitale Proprio, o “*Net Asset Value*” (NAV) in inglese, calcolato nello scenario base e il Capitale Proprio calcolato in uno dei tre scenari stressati. Gli scenari stressati sono di incremento permanente dei tassi:  $\{\Lambda = \lambda_{up}\}$ , di calo permanente dei tassi:  $\{\Lambda = \lambda_{down}\}$  e di estinzione di massa:  $\{\Lambda = \lambda_{mass}\}$ , quest’ultimo in particolare dipende dal gruppo di contratti per il quale si sta valutando il requisito. Infine, lo scenario base è  $\{\Lambda = \lambda_{BE}\}$ , che indica il tasso di estinzione atteso o “*best estimate*”. Un’importante differenza tra i tre scenari, per la valutazione dei requisiti, sta nel fatto che i primi due si riferiscono ad una modifica permanente del tasso, ovvero per tutti gli anni futuri, mentre invece l’estinzione di massa riguarda un evento catastrofe riferito ad un periodo limitato annuo, dopo il quale la frequenza di riscatto torna al valore atteso.

La formula per il calcolo dell’SCR è stata qui espressa per semplicità di rappresentazione tramite una differenza del Capitale Proprio condizionato a due diversi scenari. Quello che viene fatto più tecnicamente è il già citato *stress testing*. Sempre semplificando leggermente il ragionamento, è possibile dare un maggior dettaglio all’effettiva procedura di calcolo del requisito patrimoniale per il modulo *lapse*: in un primo momento vengono stimate, secondo una delle procedure che verranno presentate in seguito, le probabilità di recesso annue in funzione delle caratteristiche dei contratti; queste

---

<sup>39</sup>Deve valere che il risultato dell’incremento o del decremento del tasso sia compreso tra zero e uno, come si vedrà nelle formule.

stime rappresentano lo scenario atteso, la *best estimate*. Tramite queste stime vengono determinate le Riserve Matematiche dei diversi contratti, calcolando il valore attuale atteso dei flussi relativi alla polizza - che comprendono anche il possibile pagamento del valore di riscatto in caso di recesso - . L'insieme delle Riserve dei singoli assicurati costituisce le Riserve Tecniche. Sottraendo il valore delle Riserve e del Margine per il Rischio al valore degli Attivi si ottiene il NAV. Per ottenere invece il NAV nello scenario di *lapse up* si moltiplicano le probabilità base per ogni anno, fino a scadenza, per 0,5 - o meglio, si applica la formula data sopra - , e si ricalcolano le riserve e quindi il NAV; simile procedura è applicata allo scenario *lapse down*. Per quanto riguarda lo scenario *mass* si ha una leggera differenza: per il primo anno si sostituisce alla probabilità *best estimate* il valore di 0,4 o 0,7, a seconda che il contratto valutato appartenga al primo o al secondo o al terzo gruppo sopra definiti, per i restanti anni sono lasciate le *best estimate*. Come già anticipato, il ragionamento risulta semplificato, non si tengono conto ad esempio delle variazioni nel valore degli Attivi dovute al cambio di scenario, ma le conclusioni raggiunte sono le medesime. Inoltre, il discorso sembra implicare una valutazione contratto per contratto, esso però non è l'unico modo previsto dal Regolamento Delegato: è possibile fare una valutazione per gruppi di contratti che abbiano un profilo di rischio simile.

Come accennato nel capitolo precedente, l'utilizzo dei valori proposti dalla formula standard per il calcolo del requisito riferito al sotto-modulo *mass lapse* era in larga parte preferito all'adozione di uno specifico modello per la sua valutazione, dato che il livello calcolato di tale requisito era solitamente inferiore rispetto agli altri due e di conseguenza veniva escluso calcolando il massimo. Questo accadeva prima del 2021, quando l'inflazione era contenuta e i tassi di interesse di mercato erano stabili, perciò era ragionevole ottenere, sotto una certa data ipotesi predittiva della struttura dei tassi, un peso maggiore per variazioni permanenti del *lapse rate* piuttosto che per uno shock catastrofe però limitato all'immediato. Il cumulo delle minusvalenze o della riduzione delle plusvalenze sul lungo periodo era molto più importante. Con l'arrivo della pandemia e poi, successivamente, della guerra Ucraina-Russia e della crisi energetica, e ancora del conflitto israeliano-palestinese, le imprese hanno subito dei contraccolpi e l'inflazione è notevolmente incrementata, e ciò ha costretto la Banca Centrale ad aumentare i tassi e le imprese assicurative, tra cui Eurovita, si sono ritrovate gravate da forti svalutazioni negli Attivi. La situazione di crisi è stata drammatica; il numero di contratti riscattati è aumentato improvvisamente, così come le sue previsioni future, ma sorprendentemente non tanto quanto ci si sarebbe aspettati, o meglio non tanto quanto ci si potrebbe aspettare usando le stime del 40% o 70% della formula standard. Questo ha innescato una dinamica molto particolare: le assicurazioni, sebbene in grado di sopportare l'effettivo sbalzo del *lapse rate*, si sono trovate a dover allocare comunque ingente capitale poiché condizionate ancora all'uso della formula standard, impattando negativamente la situazione sui bilanci. Questo è pressappoco ciò che è accaduto al settore assicurativo tra il 2022 e il 2023. Vi sono state imprese più colpite, tra cui sicuramente Eurovita, e altre meno colpite, portafogli o singole segmenti più colpiti e altri quasi inalterati. L'impatto generale è stato importante, soprattutto dopo il crollo



di Eurovita, tuttavia non è stato tale da giustificare i valori utilizzati dalla formula standard.

In un breve articolo di inizio gennaio 2023, proveniente da un sito di informazione in ambito assicurativo, “*Intermedia Channel*”, riporta che, di fronte allo scenario di tassi alti e riscatti elevati, le imprese assicurative sono ricorse alla riassicurazione, per trasferire il rischio di estinzione di massa<sup>40 41</sup>. La scelta delle compagnie è stata dettata dalla necessità di ottenere una risposta immediata al problema, quando invece l’adozione di un modello interno avrebbe potuto richiedere molto tempo e denaro. La soluzione proposta dalla riassicurazione ha un costo per la Cedente, che può risultare sopportabile nell’immediato, ma non è detto che rimanga così nel lungo. Inoltre, le imprese riassicuratrici a loro volta devono poter valutare il *mass lapse risk*, sia per la gestione diretta che per la valutazione di prodotti che coprano il rischio *mass* trasferito dalla Cedente. Con ciò si vuole sottolineare che, sebbene la riassicurazione sia una risoluzione buona per il problema del rischio in questione, essa potrebbe esserlo solo nel breve periodo, e che in ogni caso la costruzione di un modello interno per il rischio di massa appare necessario.

Un altro scritto interessante che riporta nel dettaglio la dinamica storica che è stata prima raccontata è un documento del 27 luglio 2022 di ANIA, che presenta un approfondimento sul modulo del rischio di estinzione<sup>42</sup>. Qua sotto sono riportati due grafici che ben illustrano la portata del fenomeno *mass lapse*:

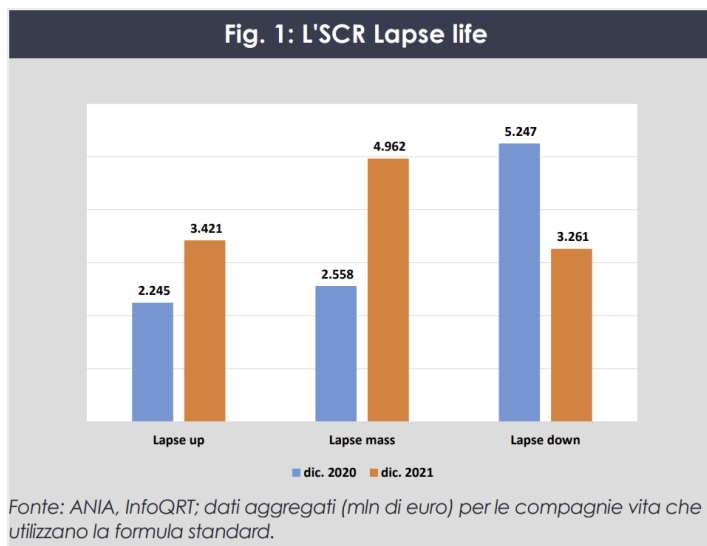


Figura 5: Dario Focarelli (2022).

<sup>40</sup>“Riccardo Sabbatini - I paradossi di Solvency II: Esopo ed il lapse risk degli assicuratori”, <https://www.intermediachannel.it/2023/01/30/i-paradossi-di-solvency-ii-esopo-ed-il-lapse-risk-degli-assicuratori/>.

<sup>41</sup>Sul tema della riassicurazione per il *mass lapse risk* si può leggere un approfondimento più dettagliato in Paul Fell (2022).

<sup>42</sup>Dario Focarelli (2022).

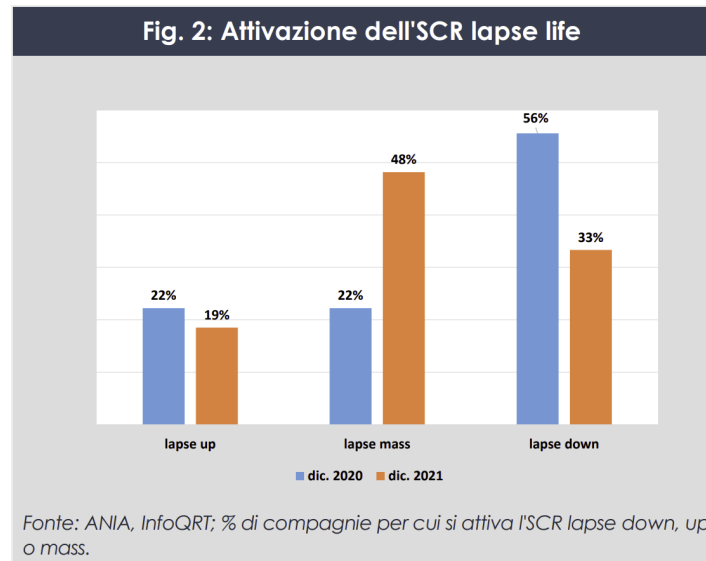


Figura 6: Dario Focarelli (2022).

Il primo presenta come il  $SCR_{mass}$  del 2022 - come dato aggregato delle compagnie - sia incrementato quasi del doppio rispetto all'anno precedente, mentre gli altri due requisiti sono calati. Il secondo invece presenta la percentuale delle compagnie per le quali il massimo, nella formula standard, risultasse proprio il  $SCR_{mass}$ , sempre un incremento doppio rispetto all'anno precedente.

L'articolo afferma inoltre:

*“Una prima analisi effettuata ha individuato [...] alcuni elementi di criticità presenti nelle attuali metodologie di calcolo della formula standard (tra cui ipotesi e parametri non coerenti con le reali dinamiche di comportamento degli assicurati)”*

In effetti risulterebbe che i valori ricavati per il rischio di estinzione di massa dagli Studi d'Impatto Quantitativi (QIS) siano solo stime approssimative causate dalla mancanza di dati a riguardo. Ciò è anche confermato in uno dei “*Calibration Papers*” di Solvency: *CEIOPS-SEC-40-10 Solvency II Calibration Paper* (2010), dove si scrive appunto: “*The empirical basis to calibrate the mass lapse event is poor*”<sup>43</sup>. Una ulteriore conferma la si può trovare in uno dei più recenti articoli riguardanti la modellizzazione del *mass lapse*, il Biagini et al. (2021), dove si può leggere: “*Interestingly, this level of 40% has not been predicted by utilizing data but is based only on expert judgment.*”. Molto interessanti in merito alla questione sono le conclusioni raggiunte dal documento tramite la loro analisi empirica: “*We thus conclude that Solvency II's mass lapse assumption corresponds to the 99.9% quantile rather than the 99.5% quantile, on which the*

<sup>43</sup> *CEIOPS-SEC-40-10 Solvency II Calibration Paper* (2010), capitolo 3, punto 3.6 “*Lapse risk*”, sotto-paragrafo 3.328

*solvency capital requirement in European insurance regulation is usually based.*” Questo però per una parte del loro campione, composto da 122 imprese tedesche, in linea con i *trend* europei, per le quali un *mass lapse rate* più realistico dovrebbe attestarsi attorno al 20 / 25%. Mentre invece per il restante campione, composto da imprese americane, la stima si attesta attorno al 50%. Questo comprova il fatto che l’uso della formula standard può portare a errori di sovrastima e di sottostima, è che vi sia pertanto una necessità di ricercare e utilizzare un modello interno specifico per tale rischio, che possa adattarsi meglio al profilo di rischio dell’impresa.

## 2.3 Le cause del rischio di estinzione e del recesso di massa

Riassumendo brevemente, l’interesse della presente Tesi è lo studio del rischio di estinzione di massa, definito come un improvviso ed inaspettato *shock* incrementativo del tasso di estinzione, similmente ad un evento catastrofico; si vuole in particolare stimare un capitale di rischio atto a coprire, con una certa probabilità, l’evento in questione. Per fare ciò è necessario indagare le cause e i fattori che influenzano il rischio e ne determinano l’impatto e la frequenza. La documentazione in proposito è molto succinta, il tema infatti è di per sé ancora poco studiato, dato che le criticità precedentemente discusse sono sorte solo negli ultimi anni. Un buon modo è partire descrivendo il semplice rischio di estinzione anticipata, per il quale esistono innumerevoli articoli e studi. Si comincerà perciò da una descrizione del comportamento degli assicurati, il “*Policyholder Behaviour*” (PHB), con riferimento al diritto di riscatto, e si passerà poi a presentare tutte le variabili per le quali si è osservato, nel corso degli anni, determinare tale comportamento, riportando anche i possibili effetti esterni che vi interagiscono e lo condizionano<sup>44</sup>.

È abbastanza consolidata l’idea di suddividere il PHB in due tipi di comportamento: uno “statico” e un “dinamico”, anche detti rispettivamente “non razionale” e “razionale” - qui la nomenclatura preferita sarà la prima, come spiegato nel paragrafo iniziale - . Si riferiscono ad una osservazione empirica della serie storica del *lapse rate*: vi è una componente di fondo abbastanza stabile nel tempo, indipendente dalle macrovariabili di mercato, la quale è solitamente attribuita a quei soggetti che decidono di recedere per motivi strettamente personali, come ad esempio problemi di liquidità; vi è poi una seconda componente che segue l’andamento di fattori economici, come i tassi di mercato contro il tasso garantito dalla polizza, e che quindi si muove nel tempo, anche con diversa volatilità. La prima è approssimativamente stazionaria ed determina un limite inferiore atteso al processo del *lapse rate*; la seconda invece è un processo non stazionario, dipendente da altre variabili, ed è detta appunto “dinamica” poiché

---

<sup>44</sup>Per una trattazione più completa del PHB si rimanda al report: *Modeling of Policyholder Behavior for Life Insurance and Annuity Products a Survey and Literature Review* (2014).

reagisce in maniera dinamica a eventi esterni e, riferendosi all'assicurato, si parla più propriamente di “*Dynamic Policyholder Behaviour*” (DPHB). Il comportamento è tanto distinto che in alcuni casi il *pool* degli assicurati si può suddividere in due gruppi riferiti alle due componenti, quando ad esempio si abbiano soggetti più o meno informati di finanza, o quando vi siano individui esperti che gestiscono per conto di terzi i contratti. Ovviamente questa distinzione si può modificare nel tempo, perché i soggetti potrebbero cambiare abitudine. Il comportamento così approssimato è in grado di spiegare la maggior parte delle serie osservate: vi sono portafogli dove la componente dinamica è assente e rimane il semplice andamento di fondo, approssimativamente stabile nel tempo, altri dove invece le variazioni sono repentine, rincorrenti il mercato.

Sono state osservate però delle differenze sostanziali nei periodi di maggior crisi: il modello così presentato subisce un *break* strutturale, più nello specifico, sembra che gli assicurati cambino il loro comportamento nel tempo. Questa osservazione nasce in primis nei modelli comportamentali finanziari, riferiti in generale a soggetti economici, come possono essere operatori di Borsa o correntisti di una banca, che si trovano colpiti da un periodo di crisi. Essi tendono a modificare il loro comportamento e a conformarlo rispetto alla massa, come una sorta di “contagio” o come un effetto a catena, il quale, dal punto di vista probabilistico, si può interpretare come una forte correlazione positiva tra i soggetti. La letteratura ha chiamato tale fenomeno “*Panic Selling*” nel caso della Borsa e “*Bank Run*” nel caso della banca. Un esempio comico, che però ben rappresenta la dinamica, si può trovare nel film “*Mary Poppins*” del 1964, quando i due fratelli protagonisti rischiano di far crollare la Banca d’Inghilterra, innescando un panico generale tra i clienti.

Questa ulteriore ipotesi rappresenta un elemento chiave per spiegare il comportamento osservato negli assicurati durante quest’ultimo periodo di tassi alti e di generale insicurezza. Condizioni simili si sono verificate anche in passato, ma solo ora si è potuto osservare pienamente il fenomeno e constatare questo effetto di correlazione tra assicurati, e ciò completa il quadro sul PHB.

Questo effetto di correlazione viene a volte chiamato “di Contagio”, per il fatto di assomigliare ad un diffondersi di una malattia; come una malattia, infatti, dipende dal numero di “malati” e dalla forza del “morbo”, e va a modificare il precedente stato di “salute” di un individuo. Per dare un esempio, preso un soggetto non interessato alla situazione patrimoniale dell’impresa presso cui è assicurato o all’evoluzione dei tassi, un assicurato “statico” si potrebbe dire, questi potrebbe cambiare comportamento venendo investito da notizie allarmanti o venendo raggiunto da più banche con proposte allettanti o anche semplicemente osservando un numero cospicuo di suoi amici o colleghi riscattare la sua stessa polizza. Presi assieme più individui, assimilabili per certe caratteristiche, ai quali è possibile associare un’unica probabilità di estinzione anticipata, questi potrebbero modificare inaspettatamente il loro comportamento e scostarsi da tale valore, per allinearsi a quello di altri assicurati con diverse caratteristiche<sup>45</sup>.

---

<sup>45</sup>Il discorso è stato semplificato, sarebbe più corretto parlare di una probabilità su un orizzonte temporale.

Riguardo ancora al *bank run*, un articolo recente sul tema *mass lapse*, che verrà presentato in un capitolo successivo, il Cheng et al. (2023), osserva come esista una sostanziale differenza tra il fenomeno del contagio nel *bank run* e nel *mass lapse*, in termini di impatto sulla compagnia. Scrive infatti:

“For deposit withdrawals, a customer liquidates deposits at the face value without paying substantial fees, that is, without forfeiture of assets. For surrenders, the payout can differ from the policy’s fair market value.”

Ovvero, mentre nel caso della banca questa deve pagare un valore di poco inferiore a quello di deposito, l’assicurazione invece paga un valore significativamente inferiore a quello del contratto, a causa delle penali di uscita. L’effetto prodotto è di una riduzione della Riserva Matematica, e quindi della passività, maggiore della riduzione degli attivi per il pagamento del valore di riscatto, con conseguente aumento del Capitale Netto<sup>46</sup>. Come già osservato, in sostanza, l’assicuratore può anche beneficiare della presenza di riscattanti, se si verificano delle giuste condizioni.

Il contagio e la correlazione sono ingredienti fondamentali per il *mass lapse*, ma di per sé non sono sufficienti, nel senso che gli assicurati possono manifestare sempre un comportamento da “copione”, identificato in letteratura con il termine “*Copycat Behaviour*”, cioè che segue la massa, ma non per questo è detto si generi la situazione catastrofale. È necessario un innesco o evento scatenante, che produca una reazione a catena di contraenti che, copiandosi, recedano l’uno dopo l’altro; un “*Panic Copycat Behaviour*”. Questo “*trigger event*” può essere spiegato tramite il raggiungimento di un certo valore o il superamento di un livello barriera di una o più variabili che condizionano il PHB, per questo è anche fondamentale analizzare i fattori da cui dipendono le tre componenti di comportamento appena descritte.

E venendo appunto alle variabili: è prima di tutto rilevante, in particolare per la componente dinamica, la presenza o no del diritto al riscatto, che rende più appetibile l’estinzione anticipata, rispetto al semplice abbandono, la cui convenienza è della sola interruzione del flusso, negativo per il contraente, dei premi, se periodici. Altra variabile che condiziona la scelta è l’“antidurata” del contratto, con cui si intende l’età della polizza ad una certa data. Infatti, si osserva, in generale, che più ci si avvicina al termine del contratto e più si riduce il numero dei riscatti osservati. A questo poi si deve aggiungere, richiamato in una nota precedente, l’assenza per legge del diritto di riscatto nei primi due o tre anni dalla stipula, per permettere la costituzione delle riserve. Altro fattore rilevante sono le già citate penali di uscita, che riducono di un certo ammontare l’importo capitalizzato all’epoca di recesso attribuito al contratto. Permettono di ridurre la probabilità che l’assicurato eserciti l’opzione.

Tutti questi fattori influenzano fortemente il DPHB, ma non solo gli unici: i più importanti sono le garanzie e i diritti contrattuali aggiuntivi discrezionali, che vengono confrontati con l’andamento dei rendimenti nei mercati finanziari. Nello scegliere se

---

<sup>46</sup>La banca non restituisce l’intero deposito poiché vengono sottratte possibili spese per commissioni o di altro genere.

recedere o meno, un assicurato dinamico osserverà quale garanzia perde, ad esempio il rendimento a scadenza ad un tasso garantito, o il bonus in caso di decesso da aggiungere al rendimento garantito base o simili. Non solo, in una Rivalutabile cercherà di stimare un possibile rendimento previsto al di sopra della garanzia di tasso base, e lo stesso sarà ancora più rilevante in prodotti come *Unit Link*, *Index Link* e nelle *Variable Annuity*, dove le garanzie sono centrali. Soprattutto, valuterà l'andamento della polizza, se in perdita o meno rispetto ai premi investiti, e recederà, anche in presenza di una forte penale, se prevederà di poter reinvestire il valore di riscatto in un altro strumento e di limitare la perdita. Perché appunto la scelta è sempre vincolata dall'andamento del mercato, da quanto è più o meno appetibile l'investimento in altri strumenti. Per questo una buona variabile utilizzata spesso nei modelli per il *lapse risk* è la differenza tra il tasso di mercato e il tasso a credito per la polizza, cioè tra quanto rende il mercato e quanto mi garantisce l'assicuratore.

Con rendimento di mercato si possono ovviamente intendere diversi flussi e quindi, in un certo qual modo, si possono avere diverse variabili: si può pensare al rendimento medio di un particolare mercato o di una certa Borsa - inteso come un portafoglio unico - , solitamente però si preferisce riferirsi ad un tasso *risk free*, come quello implicito di un titolo di stato; o ancora, ci si può riferire a strutture per scadenze di tassi a pronti o a termine, con scadenziario continuo o discreto, istantaneo o non istantaneo. Sempre per dare un esempio, un articolo scientifico, riguardante il tema del *mass lapse*, il Barsotti, Milhaud, and Salhi (2016), utilizza come variabile il tasso *forward* istantaneo. Vi è poi il problema di quando un assicurato osserva il tasso e quanto tempo impiega per decidere e successivamente recedere. Ovverosia: gli effetti di un incremento dei tassi di mercato sul tasso di estinzione anticipata potrebbero essere differiti nel tempo. Oltre al tasso osservabile di mercato, un assicurato che gestisse in autonomia i propri investimenti o un intermediario finanziario che gestisse polizze per conto di terzi, potrebbe confrontare il tasso a credito con il proprio rendimento atteso sui suoi investimenti. Un'ulteriore variabile da considerare è quindi la competenza finanziaria dei clienti, la "*Financial Literacy*", sempre di più in aumento, e similmente anche la presenza, o meno, e la numerosità, nel *pool* di assicurati, di enti che svolgano il ruolo di intermediari tra il contraente e l'assicuratore<sup>47</sup>. Quest'ultima variabile, poi, è correlata con il capitale assicurato della polizza, poiché si osserva che i soggetti privati che investono ingenti somme di denaro si rivolgono quasi sempre a intermediari che possano gestire l'investimento in maniera ottimale. Tali soggetti possono pertanto essere identificati come assicurati con comportamento dinamico: essi saranno i primi a muoversi e quelli a cui saranno pagati i valori di riscatto più alti. Per questo sono spesso raggruppati in proprie sotto-classi all'interno di una LOB, in modo da poterli valutare in modo distinto e con maggior riguardo.

A questi fattori è legato il ragionamento, tenuto nel primo paragrafo, in merito alla valutazione dell'istante ottimale di riscatto secondo i criteri matematico-finanziari. Tale

---

<sup>47</sup>Gli enti che assumono tale ruolo sono suddivisi in due categorie di istituti: "*Institutional Customers*" o "Clienti Istituzionali" e "*Institutional Investors*" o "Investitori Istituzionali".

variabile, “Istante Ottimale di Riscatto”, - molto complessa nella sua determinazione - può guidare la scelta del recedente, anche solo in termini approssimati, e l’ipotesi ad essa associata è detta in letteratura “*Moneyness Hypothesis*”, richiamando la terminologia sulle opzioni, con “*in the money option*” e “*out of the money option*”.

Questa ipotesi viene spesso confrontata con una seconda, anch’essa molto studiata e che qui è già stata citata, riguardante invece il PHB statico, detta “*Emergency Fund Hypothesis*”. Essa afferma che la polizza venga utilizzata da alcuni assicurati come fondo per accantonare risparmi, da usare per eventuali emergenze economiche varie - alcuni esempi sono già stati fatti - , nello specifico quindi per problemi di liquidità personale<sup>48</sup>. Questa ipotesi viene corroborata da alcuni studi che hanno osservato come, in alcuni portafogli, il *lapse rate* sia correlato con il tasso di disoccupazione, il quale può essere ragionevolmente visto come una *proxy* dei problemi di liquidità di un soggetto<sup>49</sup>.

Le due ipotesi appena presentate hanno come naturale conseguenza di poter associare i tipi di contratto al tipo di comportamento dell’assicurato: contratti con natura di semplice copertura saranno legati a contraenti poco proni all’abbandono, mentre prodotti di tipo risparmio vedranno un diverso ma sempre piccolo e stabile andamento del tasso di estinzione, legato all’ipotesi di fondo di emergenza, in fine prodotti con forti componenti di investimento saranno associati ad assicurati dinamici.

Tornando al PHB, si può citare un’ultima variabile che si è osservato essere legato a questo: l’età: le età più giovani sono di solito quelle più prone al risparmio e meno economicamente solide, pertanto vicine all’ipotesi di fondo di emergenza, ma sono anche quelle più reattive alle condizioni di mercato e quindi più dinamiche.

Oltre a queste variabili, proprie dell’assicurato e riguardanti nello specifico il suo comportamento, si possono osservare altri fattori che influenzano esternamente la serie dei riscatti. Un primo esempio è il tasso di mortalità del portafoglio: più aumentano i decessi e meno abbandoni si osservano, come conseguenza necessaria. Si è soliti però escludere tale variabile dai modelli perché il peso nei portafogli più dinamici - dal punto di vista dei riscatti - è irrilevante: la mortalità è maggiore nei portafogli di assicurati con età elevate o con particolari caratteristiche biologiche, quindi certamente non riguardanti prodotti di risparmio o investimento. Altro effetto esterno sono i possibili interventi dell’IVASS, il quale ha il potere di revocare o sospendere il diritto di riscatto se lo ritenesse opportuno, come è accaduto nel caso menzionato nell’introduzione di Eurovita. Un altro ancora è la stagionalità, a livello non solo mensile ma anche settimanale, dovuta a fattori come la chiusura della compagnia in specifici momenti dell’anno.

Un’altra variabile che può venir utilizzata è la dimensione dell’impresa, in termini di numero di contratti assicurati o di fatturato. È, in un certo senso, una variabile scontata e non scontata allo stesso tempo. Si potrebbe pensare che non dovrebbero esserci differenze in termini relativi, rapportando il numero di abbandoni alla dimensione, ma si è osservato, come nel caso già citato del Biagini et al. (2021), nel quale una covariata

---

<sup>48</sup>Si può citare a questo proposito l’articolo Kuo, Tsai, and Chen (2003).

<sup>49</sup>In merito a questi studi si rimanda alla nota precedente.



fondamentale è proprio questa. Al variare del volume di affari, osserva l'articolo, variano anche le numerosità relative degli abbandoni, *ceteris paribus*. Questa può essere utilizzata come una variabile approssimante di tutti quei fattori non osservabili che contraddistinguono una azienda piccola da una media e da una grande.

Sono invece più complesse da identificare le variabili che condizionano il *mass lapse*. Una prima, interessata anche dal DPHB, è legata al "rischio d'immagine", ovvero a come la compagnia stessa viene percepita dai suoi *stakeholder*, sia in senso morale - quanto l'impresa venga vista come virtuosa o meno - , sia nel senso di "*Customer Satisfaction*" e sia nel senso di solidità economica. I primi due possono essere misurati in svariati modi, utilizzando indici esterni ed interni all'impresa; in generale le aziende grandi sono già premunite di indicatori di questo tipo. Per l'ultima bisogna fare attenzione alla sua valutazione, poiché il mercato recepisce solo informazioni osservabili, rilasciate dall'impresa stessa o valutazioni fatte da esterni, come giornali d'informazione specifica; il mercato non conosce mai il reale stato di salute dell'impresa. Si possono trovare diversi indici di valutazione delle aziende facilmente reperibili, si può anche utilizzare il semplice valore delle azioni quotate, ma, trattandosi di assicurazione, un indice molto efficace proposto in letteratura è l'ultimo *solvency ratio* disponibile al mercato.

Legata al problema dell'immagine, si ha un'altra variabile importante da considerare: la capacità dell'impresa di fare "*Public Relation*", indirizzata soprattutto alla rassicurazione dei clienti in situazioni di crisi, e ad essa si può affiancare la generale capacità di gestione immediata del rischio e di risollevarlo. Com'è stato detto, l'evento *mass lapse* è improvviso, inaspettato, ma in particolar modo temporaneo e la lunghezza di questo periodo dipende in parte da quanto l'azienda è in grado di arginare l'esondazione dei recessi tramite manovre di tipo finanziario e di tipo comunicativo. Una misurazione di questo complessivo di competenze potrebbe essere già presente nell'impresa. Un fattore più evidente è la presenza o meno di crisi finanziarie o di altri eventi catastrofici o anche solo cambiamenti economici. Sebbene è possibile si verifichi un'estinzione anticipata di massa anche in assenza di tali fattori, colpendo magari una singola impresa, è ragionevole supporre però che essi siano i più decisivi e che l'evento tenda a investire più imprese nello stesso periodo. Purtroppo sono per loro natura imprevedibili e difficili da stimare nella frequenza e nell'impatto, e spesso possono presentarsi in diverse forme. Data l'assoluta incertezza che si ha su questi fattori, non si può disporre di variabili adeguate. Quello che può essere fatto piuttosto è individuare alcuni segnali che possano essere correlati con questi. La presenza o meno di conflitti potrebbe essere considerata come un elemento tale da aumentare la probabilità dell'evento *mass*. Un fattore più semplice che viene utilizzato in due modelli già citati che trattano del *mass lapse*, il Barsotti, Milhaud, and Salhi (2016) e il Cheng et al. (2023), è quello di un livello barriera, superato il quale il processo aumenta la probabilità di verificare l'evento in questione. Nel primo caso, la barriera è di tipo finanziario, ovvero si suppone che, quando la differenza tra tasso di mercato e tasso a credito supera una certa soglia  $B$ , si può verificare un incremento notevole dei recessi, il quale può a sua volta portare ad un effetto a catena per il contagio tra gli assicurati. Diversamente, nel secondo, la barriera  $B$  fissata si riferisce alla serie osservabile dei riscatti di un gruppo



di polizze pesati per il tempo. Non appena tale processo supera la soglia  $B$  si ha una certa probabilità di innescare un contagio e quindi l'evento di massa. Riassumendo, se il primo tenta di cogliere i segnali dal mercato finanziario, il secondo invece li ricerca nella situazione generale dei riscatti. Il *lapse rate* stesso, perciò, preso con un certo ritardo, può essere utilizzato come variabile influenzante il rischio di massa.

Questo approccio che fa uso della soglia permette di inserire nel modello quello che è stato definito all'inizio del paragrafo come il "*trigger event*". La stima della barriera  $B$  può essere molto complessa. I due articoli ad esempio, nei loro rispettivi approcci, propongono di basarsi sulla valutazione di personale esperto nel tema dei riscatti. Il metodo è il cosiddetto "*Expert-Based Judgment Method*", il quale è fortemente soggettivo e difficilmente giustificabile.

Per quanto riguarda l'effetto di contagio, anch'esso, come visto, un fattore legato al *mass lapse*, nell'ambito dei recessi non esistono attualmente misure o variabili studiate in proposito, che possano permettere di dare una valutazione sulla repentinità del fenomeno, sul suo impatto economico e su quanto sia suscettibile rispetto ad eventi esterni o se si modifichi nel tempo. Lo studio del fenomeno è attualmente molto nuovo in ambito economico-finanziario e deve ancora essere tradotto nei problemi assicurativi per il PHB, soprattutto perché si potrebbero osservare delle differenze sostanziali tra il comportamento dei soggetti nei mercati finanziari e il comportamento degli assicurati, in caso di contagio.

Si conclude qui il sunto di quanto la letteratura scientifica afferma riguardo al comportamento degli assicurati che recedono e di quali sono i fenomeni e le variabili che lo condizionano. Questo capitolo però non esaurisce la discussione, non solo per la lunghezza dell'argomento, ma anche e soprattutto perché è necessario affrontare, e verrà affrontato nel prossimo capitolo, la problematica della modellizzazione di tale fenomeno, per il semplice rischio di estinzione e per rischio di massa; quali sono i modelli che sono stati proposti e ora rigettati, quali sono le criticità che rendono difficile la sua costruzione e stima, e cosa dice la letteratura attuale, molto minuta per quanto riguarda il *mass lapse*, in merito.

### 3 La modellizzazione del *lapse rate* e del *mass lapse risk*

Già nel capitolo precedente si potevano cogliere le criticità che rendono particolarmente ostica la modellizzazione del *lapse rate* e soprattutto dell'abbandono di massa. Con questo capitolo si vuole andare ora a esplicitare tutte queste criticità nel dettaglio, affrontandole attraverso le descrizioni delle metodologie utilizzate per il modellamento del tasso di recesso. In particolare il primo paragrafo affronterà il tema della modellizzazione del *lapse rate*, mentre con il secondo si andrà a spiegare i motivi che rendono tali approcci non consoni per la valutazione del rischio di massa e si andrà conseguentemente a presentare quattro documenti proponenti quattro approcci nuovi, specifici per il *mass lapse*.

#### 3.1 L'approccio ordinario alla modellizzazione del *lapse rate*

Per quanto complesso sia il fenomeno, gli approcci si rifanno in realtà per la maggior parte a modelli classici della probabilità, sia perché sono più immediati da implementare, sia perché, applicati alla serie dei tassi, permettono di ottenere in generale buoni risultati<sup>50</sup>. Visto tutto quanto è stato detto nel capitolo precedente sul tasso di riscatto, esso può essere trattato sotto due strade, una che fa ricorso ai metodi della tecnica attuariale Vita e l'altra invece che segue la tecnica attuariale Danni, similmente a quanto accade per il ramo Infortunio e Malattia, che si trova a metà tra i due rami principali. Si potrà vedere che in realtà gli approcci possibili per l'analisi del *lapse rate* sono interconnessi.

##### 3.1.1 I metodi della tecnica attuariale Vita per l'analisi del *lapse rate*

È possibile farlo rientrare nell'ambito delle tecniche attuariali Vita se si va a costruire il modello a partire dalla singola polizza, e definendo per questa il numero aleatorio “*durata della polizza*”. Il contratto, infatti, può terminare per raggiungimento della scadenza, per decesso o per recesso, con e senza riscatto<sup>51</sup>. Si possono quindi, a partire

---

<sup>50</sup>Con “modelli classici” si intende l'insieme delle tecniche diventate ormai comuni nella modellizzazione di fenomeni aleatori, come i modelli GLM, i modelli per più cause d'uscita, i processi di *Poisson*, etc. Quindi, in particolare, per la modellizzazione del rischio di riscatto non si è ritenuto necessario ricorrere a tecniche più desuete o costruite ad hoc per il *lapse rate*.

<sup>51</sup>Si potrebbero inserire tutte le possibili modifiche contrattuali già elencate o altri possibili motivi di uscita, di tipo legale ad esempio; per il discorso attuale è sufficiente la semplificazione.

da un dataset adeguato, stimare le probabilità di estinzione anticipata utilizzando i modelli per più cause di uscita<sup>52</sup>. Gli assicurati devono potersi considerare omogenei, di conseguenza bisogna tener conto per lo meno di età del soggetto, anno di sottoscrizione, antidurata e capitale assicurato. I dati potrebbero essere censurati - quasi sicuramente considerate le durate dei contratti vita - , richiederebbero perciò l'utilizzo di stimatori come quello di *Kaplan-Meier*<sup>53</sup>. Il prodotto del processo sono le tavole di sopravvivenza, per la mortalità e soprattutto per i recessi.

Ovviamente le basi statistiche richieste sono enormi e il dispendio di tempo per far funzionare la procedura è elevato. Di solito per la mortalità il problema è più agevole perché più studiato anche al di fuori del mondo assicurativo e si possono trovare innumerevoli *database* pubblici di alta qualità. Mentre, per quanto riguarda il numero di riscatti osservati, i dati sono difficili da reperire poiché le imprese non li rendono pubblici, dato che vengono giudicati come dati sensibili sullo stato della compagnia<sup>54</sup>. Il problema principale tuttavia sta nel fatto che è ragionevolmente molto difficile supporre una buona omogeneità tra le classi: mentre i decessi sono un fenomeno principalmente legato alla biologia, per il quale i miglioramenti e i peggioramenti si distribuiscono abbastanza equamente su tutta una popolazione e si può osservare una tendenziale crescita stabile, quindi più prevedibile, nelle durate di vita, il *lapse rate* dipende invece da innumerevoli fattori che hanno impatti rilevanti, anche solo prendendo a modello una singola compagnia: questa può osservare su diversi portafogli tassi assolutamente variegati, non solo nello spazio ma anche nel tempo, con rotture possibili delle serie; tale è di fatto il problema di lavorare con dati panel di tipo economico, ma se ne discuterà meglio più avanti. Ad esempio, nell'analisi statistica successiva, il campione, formato da tre portafogli riferiti a tre gruppi di rischio omogeneo di una medesima tipologia di contratti, presenta tre tassi di riscatto nettamente distinti.

Il comportamento del tasso di estinzione è, in sostanza, più mobile e imprevedibile rispetto alla mortalità. Per questo motivo è necessario limitarsi a considerazioni temporalmente brevi o a introdurre ulteriori variabili nel *dataset* e aggiungere all'analisi un passaggio ai modelli di regressione, in particolare ai GLM, dato che la variabile dipendente ammette valori solo compresi tra zero e uno. Si possono implementare GLM per dati *cross section* o per dati *panel*, di tipo regressione binomiale con funzione di legame canonico logit o altri legami, tipo probit o cloglog, introducendo il maggior numero di variabili significative<sup>55</sup>. Un buon criterio per individuare ex ante dei regressori da introdurre nella matrice *design* può essere quello di farsi guidare dalle ipotesi di tasso e di fondo di emergenza, sopra descritte, e sottoporle a test di significatività.

Questo tipo di approccio è molto utilizzato e si può citare a questo proposito un ar-

---

<sup>52</sup>Sui modelli per più cause di uscita si rimanda al libro Dickson, Hardy, and Waters (2019).

<sup>53</sup>Per una trattazione sullo stimatore di *Kaplan-Meier* si rimanda all'articolo Goel, Khanna, and Kishore (2010).

<sup>54</sup>Per ottenere delle basi di studio per la mortalità basta far riferimento agli enti pubblici di statistica, presenti in ogni stato, in particolare in Europa si ha l'EUROSTAT che fa capo a tutti gli istituti degli stati membri, tra i quali per l'Italia è l'ISTAT.

<sup>55</sup>Sui Modelli Lineari Generalizzati (GLM) si faccia riferimento al libro: Dobson and Barnett (2018).

titolo recente comprensivo sia dello studio tramite modelli per più cause di eliminazione sia dei GLM, il Milhaud and Dutang (2018).

Un altro approccio classico dei modelli sulla mortalità è il “*Lee-Carter*”, il quale permette di tener conto dell’evoluzione temporale del fenomeno<sup>56</sup>. Di base il modello considera come soli fattori l’età e l’anno, ma può essere aumentato aggiungendo effetti di coorte. Una sua applicazione al *lapse rate* potrebbe sostituire l’età anagrafica con l’età della polizza, cioè l’antidurata, ma anche così avrebbe ben poco significato, dato che il *Lee-Carter* fa riferimento, nella sua struttura, agli studi fatti sulla mortalità in sé, che, come detto, si comporta in maniera molto diversa dal recesso; sarebbe quindi richiesto di trovare un’altra forma funzionale. Inoltre, sarebbe necessario introdurre anche tutte le variabili già discusse, sicuramente la differenza tasso di mercato-tasso accreditato, e tener conto del comportamento più o meno dinamico degli assicurati. Non sono state trovate in letteratura proposte simili.

### 3.1.2 I metodi della tecnica attuariale Danni per l’analisi del *lapse rate*

Lasciando la strada della tecnica vita, si può adesso imboccare quella della tecnica Danni. La stessa problematica dell’analisi del *lapse rate* può essere inquadrata in quest’ambito, nello specifico ricorrendo ai metodi “*Frequency-Severity Based*”. Infatti, si possono considerare i diversi portafogli dell’impresa come entità soggette all’evento sinistro: “*danno, in termini monetari, causato dalla perdita di un assicurato*”. Ovvero, preso un portafoglio di soggetti omogenei, che in particolare hanno contratto la polizza alla stessa data e con il medesimo termine, osservando la sua evoluzione nel corso degli anni, esso sarà colpito da un certo numero di uscite per recesso del contratto, le quali causeranno degli esborsi di diverso ammontare per i riscatti, e ciò per i diversi anni per i quali la polizza rimane in essere. È possibile ottenere il costo cumulato fino all’istante  $t$  come somma dei singoli costi dovuti al pagamento del valore di riscatto per il numero di soggetti usciti in tale arco di tempo. Supponendo di poter far rientrare il problema all’interno della *Teoria Collettiva del Rischio*, sotto le ipotesi base del modello, il processo così definito avrebbe legge *Poisson Composto*<sup>57</sup>; stimandone i parametri è possibile simulare il processo per ottenere il quantile desiderato della distribuzione, fissato un dato intervallo temporale.

L’approccio si concentra molto sulla valutazione del danno economico, quando piuttosto, seguendo la Direttiva, si preferisce solitamente focalizzarsi sul solo studio della frequenza per poter stimare la probabilità di riscatto e utilizzarla in un successivo momento per la valutazione dell’impatto. Vi è però una problematica ancora più rilevante nell’utilizzo della Teoria: le ipotesi riguardanti il processo di conta e il suo legame con il processo del costo per i riscatti. Non si può supporre che i riscatti abbiano una distribuzione uniforme nel tempo, come invece implicherebbe l’utilizzo di processi di *Poisson Omogenei*, e che siano indipendenti da altre variabili, come l’antidurata e i tassi di interesse, da cui potrebbe dipendere anche il costo per riscatto. Si pensi ad esempio al

---

<sup>56</sup>Per il modello *Lee-Carter* si rimanda all’articolo Girosi and King (2007)

<sup>57</sup>In merito al tema della *Teoria Collettiva del Rischio* si può fare riferimento a Kaas et al. (2008).

fatto che i soggetti dinamici, i quali tendenzialmente hanno capitali assicurati elevati, e quindi valori di riscatto alti, in periodi di tassi alti tenderanno a recedere, producendo una correlazione tra numero di abbandoni osservati e valori di riscatto pagati. Un ulteriore problema sta anche nel fatto che il *lapse rate* è dipendente dal numero degli individui presenti ad una data epoca, mentre il modello potrebbe portare ad un numero di arrivi superiore ai possibili. Il problema è irrilevante se si considera un portafoglio aperto, per il quale le uscite non raggiungono mai soglie di totale estinzione, è più grave di contro nel caso di un *pool* chiuso, ma per ambo i casi è possibile imporre un limite superiore prendendo il numero di arrivi in un certo anno e prendendo il minore tra quello e il numero di assicurati presenti.

Gli aggiustamenti che possono essere fatti per i precedenti problemi elencati sono ad esempio di passare a processi di *Poisson Non Omogenei*, per i quali si definisce una funzione di intensità che dipende dal tempo e, nel caso di portafoglio chiuso a nuovi ingressi, anche dall'antidurata<sup>58</sup>. Anche in questo caso vi sono due problemi: il primo risiede nel fatto che il tasso di riscatto dipende anche da altri fattori, il secondo invece che la funzione è deterministica, mentre è più realistico attendersi un processo stocastico, che permette di avere una miglior valutazione dell'errore. Si può passare allora a processi più complessi ma più completi: i *Processi di Cox*, che prevedono come parametro del processo di *Poisson* un processo stocastico scelto<sup>59</sup>. Discretizzando l'intensità, la si può ricondurre a modelli di regressione generalizzati, mantenendola continua si passa all'uso di processi a parametro continuo. Un esempio già citato di articolo che ne fa ricorso è il Barsotti, Milhaud, and Salhi (2016), che utilizza una modifica dei processi di *Hawkes*. Un secondo documento da citare è il Cheng et al. (2023), che non solo utilizza i processi di *Hawkes*, ma si preoccupa di modellizzare anche l'impatto economico dovuto ai riscatti sul patrimonio, nell'ottica *Frequency-Severity based*. I due articoli verranno meglio presentati successivamente.

Altra strada per far fronte alle critiche presentate potrebbe essere quella di utilizzare metodi simili al *Chain Ladder*, preferibilmente stocastici<sup>60</sup>. Preso un gruppo di contratti visti in una fissata epoca, questo può essere rappresentato come un insieme di coorti con diversa antidurata. Il fine è quello stimare in ottica di *run-off* il *lapse rate* futuro tramite le osservazioni passate e opportune ipotesi sulla serie, come un modello moltiplicativo a due o più fattori. Non sono stati trovati esempi di applicazione della procedura o di una sua variante, ma tale approccio può essere ricondotto, con gli opportuni accorgimenti, ai GLM, che sono in generale più versatili<sup>61</sup>; per questo motivo probabilmente non sono stati proposti in letteratura modelli di questo tipo.

A chiusura del discorso sui modelli tipo Danni, si può aggiungere che la lunga durata dei contratti e il fatto che l'evento per singolo assicurato sia irripetibile, distanziano tale

---

<sup>58</sup>Sui modelli di *Poisson* non omogenei si rimanda al libro: Brémaud (2020).

<sup>59</sup>Sulla teoria in merito ai processi di *Cox* si faccia riferimento sempre al libro: Brémaud (2020).

<sup>60</sup>In merito al *Chain Ladder* o *Metodo della Catena* si faccia riferimento al libro: Wüthrich and Merz (2008).

<sup>61</sup>Per una dimostrazione del legame tra GLM e *Chain Ladder* si veda il Lemma 2.16 del libro, citato alla nota precedente: Wüthrich and Merz (2008).

fenomeno da quelli tipici dell'ambito Danni, mentre lo avvicinano di più a prodotti che hanno ad esempio una copertura per l'invalidità permanente. È interessante notare, in merito alla distinzione tra approcci della tecnica Vita e approcci della tecnica Danni per l'analisi del *lapse rate*, come anche il Regolamento Delegato predisponga due metodi di calcolo per il rischio di estinzione anticipata malattia, uno simile al modulo *Life* (SLT) e uno invece simile al non *Life* (NSLT)<sup>62</sup>.

### 3.1.3 Gli alberi di Regressione per l'analisi del *lapse rate*

Un documento del 2018 scritto per il “*SOA Asia Pacific Annual Symposium*” dalla “*Society of Actuaries*” (SOA)<sup>63</sup>, presenta in due diapositive una raccolta dei modelli presentati qui precedentemente ed altri modelli sul *dynamic lapse*, suddivisi in due gruppi, per distinguere quelli che si basano su variabili interne, come le caratteristiche del portafoglio, e quelli che si basano su variabili esterne, come i tassi *risk free* e la disoccupazione.

Panel A: Focus on environmental variables							
Authors	Dar and Dodds (1989)	Outreville (1990)	Kuo et al. (2003)	Kim (2005a)	Kim (2005b)	Cox and Lin (2006)	Kiesenbauer (2011)
Country	UK	US, Canada	US	South Korea	US		Germany
Lapse Data Provider	Industrial Life Offices' Association (ILOA)	American Council of Life Insurance (ACLI); Life Insurance and Market Research Association (LIMRA)	American Council of Life Insurance (ACLI)	One Korean company	Society of Actuaries' Risk Management Task Force (data on one US company)	Society of Actuaries' Risk Management Task Force (data on one US company)	German regulator BAFin and self-collected data
Type of Policy	Endowment policies	Whole-life and ordinary life insurance	All ordinary life insurance contracts*	Endowment, annuity, protection plan, education	Single premium deferred annuities	Single premium deferred annuities	Endowment, annuities including disability and long-term care, term life, group business, other 1997-2009
Time Period	1952-1985	1955-1979 (annually), 1980-1998 (semiannually)	1951-1998	1997-2000	1993-2003	1993-2003	
Methods	Standard time series regressions	Standard time series regressions	Cointegration techniques	Logit and complementary log-log model	Logit	Tobit	Logistic regression
Variables for EFH and IRH	EFH: growth in the level of unemployment and the level of actual unemployment relative to time trend IRH: internal and external rate of return	EFH: yearly unemployment rate IRH: real rate of return on alternative assets	EFH: yearly unemployment rate IRH: market interest rates	EFH: Unemployment rate IRH: reference market rate - crediting rate - surrender charges Policy age, GDP growth, surrender charge	EFH: Monthly annualized US unemployment rate IRH: Annualized five-year Treasury bond rate minus the policy credited rate Policy age, GDP growth, surrender charge	EFH: Monthly annualized US unemployment rate IRH: Annualized five-year Treasury bond rate minus the policy credited rate Policy age, GDP growth, surrender charge	EFH: Unemployment rate IRH: Bond Performance, Stock Performance Buyer confidence, GDP, company age, distribution, legal form, company size, participation rate
Other variables	/	Real transitory income per capita, price of insurance, anticipated inflation rate.	/				
Main Results	EFH +; IRH -	EFH +	EFH -, IRH +	Lapse depends on additional exogenous factors beyond interest rates and unemployment rates	Logit Model closely fit the experience of the data even under extreme financial conditions	Tobit model is better than logit model; poisson and the negative binomial regression model are more appropriate to model lapse	Factors beyond interest rates and unemployment influence lapse behavior, including company characteristics

Notes: EFH: Emergency fund hypothesis; IRH: Interest rate hypothesis; \*: The covered policies include universal life, variable life, variable-universal life, and traditional whole life that includes limited payment, continuous premium, joint whole life, single premium, adjustable life, monthly debit, and other fixed-premium products), term insurance (including decreasing, level and other term, term additions), and endowment insurance.

Figura 7: Una raccolta dei modelli per il *dynamic lapse* per variabili di portafoglio. Fonte: D. Kim (2018).

La tabella 7 e la tabella 8 riassumono, oltre agli articoli e le rispettive tecniche utilizzate per la modellizzazione, anche i risultati sperimentali ottenuti dai diversi autori<sup>64</sup>. In parte, tali risultati, sono stati riassunti nel precedente capitolo.

<sup>62</sup>REGOLAMENTO DELEGATO (UE) 2015/35 DELLA COMMISSIONE (2014).

<sup>63</sup>D. Kim (2018).

<sup>64</sup>Per completezza vengono messi di seguito tutti gli articoli citati nelle due diapositive: Dar and Dodds (1989), Outreville (1990), Kuo, Tsai, and Chen (2003), C. Kim (2005), Cox and Lin (2006), Kiesenbauer (2012), Renshaw and Haberman (1986), Kagraoka (2005), Cerchiara et al. (2009), Milhaud, Loisel, and Maume-Deschamps (2011), Eling and Kiesenbauer (2014).

Panel B: Focus on product and policyholder characteristics					
Authors	Renshaw and Haberman (1986)	Kagraoka (2005)	Cerchiara et al. (2009)	Milhaud et al. (2010)	Eling and Kiesenbauer (2011)
Country	Scotland	Japan	Italy	Spain	Germany
Data Provider	Pool of 7 companies	1 company	1 company	1 company	1 company
Type of Policy	Endowment Whole-life Temporary insurance	Annuity-type personal accident	Saving	Endowment	Endowment Annuity Term life
Time Period	1976	1993-2001	1991-2007	1999-2007	2000-2010
Methods	Logistic regression model Binomial model	Poisson model Negative binomial model	Poisson model	Logistic regression model Classification and regression tree model	Binomial model Negative binomial model
Variables	Policyholder age and sex Contract age Product type Company	Policyholder age and sex Seasonality Unemployment rate Heterogeneity	Policyholder age Contract age Product type Calendar year	Policyholder age Contract age Product type Sum insured Risk premium Saving premium	Policyholder age and sex Contract age Product type
Main Results	four factors identified as important (age at entry, duration of policy, office, type of policy); additionally, significant interaction between policy type and duration	surrender of the insurance contracts is explained by change of unemployment rates and time elapsed from contract date.	Shows the importance of policy duration, but also show the sensitivity of lapse rates to calendar year of exposure, to product class and policyholder age	Find that duration and profit benefit option are essential	Product characteristics such as product type or contract age and policyholder characteristics such as age or gender are important lapse drivers

Figura 8: Una raccolta dei modelli per il *dynamic lapse* per variabili di portafoglio. Fonte: D. Kim (2018).

Tra i modelli elencati, si possono notare due non ancora discussi: gli “*Alberi di Regressione e Classificazione*”, o CART, e i modelli per le serie storiche<sup>65</sup>. I primi vengono utilizzati per poter suddividere il campione di individui utilizzando un insieme di variabili. Applicato allo studio del comportamento dell’assicurato, permette di individuare quelli che sono i fattori che contraddistinguono i soggetti con il più alto rischio di recesso. Come fa notare però l’articolo riportato nella tabella, il Milhaud, Loisel, and Maume-Deschamps (2011), questo metodo non permette di stimare la probabilità di riscattare, ma solo di appartenere ad una data classe, per questo utilizza il metodo in congiunta con la regressione logistica: dal CART seleziona le variabili rilevanti e le utilizza poi nella regressione per stimare al meglio la probabilità di recedere. La procedura quindi non dev’essere utilizzata da sola e serve a classificare il *pool* degli assicurati e a individuarne i fattori di rischio che caratterizzano la variabile dipendente.

### 3.1.4 L’analisi del *lapse rate* come serie storica

Una critica che si può muovere a questa procedura e in generale anche ai LM ordinari sta nel fatto di non considerare l’evoluzione temporale - che è una variabile distinta dall’antidurata - , cosa che invece può essere implementata tramite l’altra tipologia di modelli riportati in figura e non è ancora discussi in questa sede, i modelli per le serie storiche. Essi prendono la forma tipicamente di modelli di regressione lineari, i

<sup>65</sup>Nei due panel si possono trovare anche altri due approcci che possono sembrare non ancora discussi: i modelli “*Tobit*” e le tecniche di “*cointegrazione*”. In realtà i primi fanno parte dei modelli di regressione, applicati però a dati censurati, i secondi invece rientrano nei modelli per le serie storiche, come verrà spiegato più avanti.



quali permettono però in aggiunta di introdurre l'effetto temporale, espresso tramite una possibile correlazione seriale del processo stesso ed anche, se presenti altre variabili temporali, in una possibile correlazione con diversi ritardi di queste. Il fenomeno della correlazione seriale, o autocorrelazione, fa cadere l'ipotesi di indipendenza delle unità statistiche degli LM ed è una tendenza che si osserva spesso nell'evoluzione temporale di serie storiche economiche. L'ipotesi di indipendenza è fondamentale per far sì che gli stimatori OLS siano consistenti e asintoticamente normali, ovvero che convergano al valore reale del parametro stimato e che abbiano distribuzione limite *Normale*. I modelli per le serie storiche permettono di introdurre la dipendenza delle unità sotto forma di una relazione lineare con i valori assunti dalla serie negli istanti precedenti. È necessario però che questa relazione scompaia dopo un certo ritardo per avere le stesse proprietà dello stimatore OLS. È richiesto inoltre che la variabile dipendente e i regressori siano congiuntamente stazionari, ipotesi che fa da controparte all'ipotesi d'identica distribuzione per dati *cross-section*.

Per dare un esempio pratico, riferito alla serie dei tassi di recesso osservati lungo un periodo, se il modello scelto fosse una regressione logistica, per semplicità con un'unica covariata, il tasso di mercato contro il tasso accreditato, esso avrebbe stime corrette solo se le due variabili fossero, congiuntamente, indipendenti e identicamente distribuite, cosa assolutamente non scontata, in particolare per la covariata e per l'errore del modello<sup>66</sup>. Infatti, questa specifica covariata è naturalmente pensarla come una serie autocorrelata e, inoltre, il tasso di riscatto, in periodi di crisi, potrebbe manifestare una autocorrelazione, come osservato nel paragrafo precedente. Non solo, si può supporre ragionevolmente che esista una possibile correlazione - e quindi dipendenza - del tasso di recesso con i ritardi della covariata, dato che un assicurato osserva prima il tasso di mercato e poi successivamente decide di recedere, e il suo recesso potrebbe avvenire o essere registrato con un certo ritardo<sup>67</sup>.

Rimanendo nell'ambito delle serie storiche, i modelli a cui si fa riferimento classicamente sono gli "*Autoregressivi di ordine p*", o AR in breve, e quelli "*a Media Mobile di ordine q*", o MA, i quali poi possono essere riuniti in un unico processo, che prende il nome dalla fusione delle abbreviazioni dei due, ovvero gli ARMA, di ordine  $p$  e  $q$ . Si parla anche di modelli ARIMA, con cui ci si riferisce ad un processo ARMA "*integrato di ordine d*", la serie cioè viene "integrata" o differenziata  $d$  volte. Questi vengono utilizzati per i processi non stazionarie poiché la differenziazione può eliminare, in alcuni casi specifici, la non stazionarietà<sup>68</sup>.

Soffermandosi in particolare sugli AR e i MA, l'interpretazione dei primi è di un processo dipendente da uno o più valori passati della serie, più una variazione casuale, i secondi invece rappresentano una serie che subisce degli *shock* casuali, il cui impatto permane fino ad un certo ritardo. L'idea della dipendenza temporale della serie con

---

<sup>66</sup>Con stime corrette si intende di consistenti e asintoticamente normali.

<sup>67</sup>Del problema del ritardo nella registrazione di un recesso se ne accennerà più avanti.

<sup>68</sup>Per una definizione più completa dei modelli AR e della stazionarietà si vedano le Appendici, mentre per una trattazione sui processi ARMA si faccia riferimento al libro: Brockwell and Davis (2002).



i valori assunti precedentemente si adatta bene al problema della correlazione e del contagio che si possono manifestare nei riscatti osservati. È però un fenomeno legato solamente al rischio del riscatto di massa, come discusso in precedenza. In effetti gli articoli che studiano il *lapse rate* come serie storica non fanno uso né degli AR né dei MA, dato che appunto il fenomeno del *mass lapse* è ed era ancora poco osservato e studiato, essi si rifanno piuttosto a modelli di tipo regressione multivariata per serie temporali, ovvero che dipendono solo da una o più variabili esogene, nei loro possibili ritardi. Un esempio a metà tra le due strade, che implica l'uso di modelli autoregressivi ma che studia la sola relazione tra variabile dipendente e covariate, è il Kuo, Tsai, and Chen (2003), il quale utilizza le tecniche di “*cointegrazione*” per verificare le ipotesi di tasso d'interesse e di fondo di emergenza. La cointegrazione permette di studiare la relazione tra due serie integrate di ordine uno, soggette a un *trend* stocastico comune o simile, come un  $AR(1)$  con parametro unitario, utilizzando un fattore di correzione, l'“*Error Correction Term*”.

Un problema legato all'uso dei modelli ARMA per il *lapse rate* sta anche nel fatto che quest'ultimo assume valori tra zero e uno, il supporto della variabile è cioè limitato, mentre invece gli ARMA ipotizzano che la variabile dipendente possa assumere valori reali. Per superare l'ostacolo si può ad esempio trasformare la variabile tramite la curva logistica o utilizzando una qualsiasi funzione di ripartizione continua.

Sempre vedendo il *lapse rate* in ottica di serie storica, si possono aggiungere altre considerazioni sul possibile comportamento del tasso. Innanzitutto, questo potrebbe presentare un *trend*, oltre ad effetti di ciclicità e stagionalità. Il loro studio è fondamentale non solo perché la conoscenza di essi porta informazioni aggiuntive sulla serie stessa, ma anche e soprattutto perché è sempre necessario nei modelli poterle eliminare per avere una reale comprensione dell'andamento di fondo della serie, che si esprime poi nell'applicazione del giusto modello e per la stimabilità dello stesso. Se ad esempio la serie presentasse un *trend*, vorrebbe dire in particolare che la serie non è stazionaria e pertanto mancherebbe di un'ipotesi fondamentale per la stima dei parametri dei processi ARMA. È chiaro poi che per poter fare previsioni è necessario conoscere questi tre fattori. Sono interessanti per l'analisi del tasso di estinzione anticipata soprattutto la stagionalità e la ciclicità: la prima perché in periodi di festività, come nelle prime settimane dell'anno, ci si possono aspettare meno recessi, la seconda invece perché la serie dipende, in una certa misura, da quantità economiche - date per vere le ipotesi più volte citate di tasso e di fondo di emergenza - , le quali sono spesso soggette a cicli economici, come il rendimento medio nei periodi di mercato in crescita, il “*Bull Market*”, e nei periodi di mercato in decrescita, il “*Bear Market*”; tale effetto poi lo si può ragionevolmente associare al rischio di estinzione di massa, dato che è in periodi di congiuntura che si possono manifestare *shock* estremi. Il *trend* di contro è meno interessante, a causa del comportamento dinamico degli assicurati: in un portafoglio poco dinamico la serie, come descritto nel capitolo precedente, rimane stabile su un certo tasso minimo, mentre in un portafoglio dinamico i movimenti seguiranno i tassi, che presentano un *trend* solitamente stocastico, difficile da prevedere. Un articolo che si può citare, molto recente, sul tema della modellizzazione del *lapse* come serie storica,

che ne analizza il *trend* e il ciclo - assunti deterministici - è il Koijen, Lee, and Van Nieuwerburgh (2024).

Un ulteriore tema, di fondamentale importanza nell'analisi di una serie storica, e interessante il tema del *lapse rate*, è quello del “*Break Strutturale*”, ovvero quando il processo sottostante si modifica nella sua struttura a causa di un evento esterno non considerato nel modello<sup>69</sup>. Questo si traduce ad esempio in una temporanea o permanente modifica del valore dei parametri; se il modello venisse stimato senza considerare il *break*, si otterrebbero stime errate, dato che di fatto il fenomeno porta alla non stazionarietà<sup>70</sup>. L'abbandono di massa può essere visto come un effettivo *break* strutturale, nel senso che il comportamento degli assicurati si modifica improvvisamente a causa di un evento esterno. Per spiegare meglio: un modello, precedentemente calibrato sulle informazioni disponibili, prive però di dati riguardanti abbandoni di massa, si dimostra inadeguato a cogliere la rottura. Se l'intento è quello di modellizzare solo il comportamento ordinario, bisogna allora poter separare la serie in due parti per eliminare il *break*, se l'intento – com'è il caso della presente ricerca – è piuttosto quello di integrare questi eventi di rottura, allora è necessario modificare la struttura stessa del modello.

Si potrebbero verificare *break* strutturali con un effetto permanente sulla struttura, seguendo però quanto affermato nel *Calibration Paper* di *Solvency* citato in precedenza, il *mass lapse* lo si deve ipotizzare come un evento dalla durata temporanea<sup>71</sup>. Di conseguenza esso deve essere studiato come un *break* strutturale temporaneo, un salto nella serie, dopo il quale il tasso ritorna ai valori precedenti<sup>72</sup>. Si può pensare ad esempio ad una modifica della legislazione in materia di riscatto, o a modifiche repentine nella consistenza del portafoglio, come un afflusso di assicurati dinamici, o ancora un incremento immediato dei tassi di rendimento ad opera della Banca Centrale, o al crollo di qualche banca o assicurazione importante - come il caso di Eurovita -, tale da far sembrare il settore poco sicuro, o in ultimo una modifica nella procedura di richiesta di recesso dal contratto, che permetta di renderla più immediata o più difficile; si pensi a quanto avvenuto storicamente, dalla possibilità di recedere solo andando in presenza in sede, all'aggiunta dell'opzione di richiesta tramite telefono e poi ancora tramite sito *web* ed attualmente tramite applicazione su telefono.

---

<sup>69</sup>Per essere più precisi, non si modifica il comportamento della serie, ma è piuttosto il modello associatole che non è in grado di cogliere gli effetti della variabile esogena che ha causato il *break*.

<sup>70</sup>Come ad una nota precedente, con stime corrette si intende non consistenti e asintoticamente *Normali*.

<sup>71</sup>*CEIOPS-SEC-40-10 Solvency II Calibration Paper* (2010)

<sup>72</sup>Per chiarezza: anche un incremento improvviso permanente della serie dei tassi potrebbe essere visto come “catastrofale”, ma dal punto di vista dalla normativa questo rischio è già colto dal sottomodulo per il rischio di incremento permanente del tasso di estinzione. Se nella modellizzazione del *mass lapse* si andasse a considerare anche questo tipo di fenomeno, si produrrebbe un doppio accantonamento per lo stesso rischio.

### 3.1.5 La diverse definizioni del tasso di riscatto

L'ultimo punto del precedente paragrafo, già richiamato in precedenza, è legato al problema della definizione di *lapse rate*, nello specifico: *quali flussi utilizzare per calcolare il tasso?* Che nel caso della richiesta di riscatto si traduce in: *si usa il numero di richieste di riscatto o i riscatti liquidati?* Né la Direttiva né il Regolamento Delegato, o gli Studi Quantitativi, fanno menzione di una specifica formula da utilizzare. È lasciato quindi un certo livello di discrezionalità nella scelta delle quantità, purché il tasso ottenuto sia sensato nel rappresentare il fenomeno dei recessi. La formula utilizzata caratterizza in maniera peculiare la serie ottenuta e di conseguenza condiziona la scelta del modello più adeguato. Vi è un problema simile per quanto riguarda il tasso di mortalità, con riferimento alla costruzione del tasso a partire dai decessi, ma, data l'abbondanza di studio sul fenomeno, la nomenclatura in letteratura si è standardizzata, fissando ad esempio le definizioni di “*Tasso di Mortalità*” e di “*Tasso di Mortalità Centrale*”, quest'ultimo più spesso utilizzato nella modellistica<sup>73</sup>. Di seguito vengono presentate alcune possibili definizioni per il tasso di estinzione:

Si consideri un portafoglio, chiuso a nuovi ingressi, di  $n$  contratti assicurativi, stipulati alla stessa epoca e con la stessa durata, indicati rispettivamente con 0 e  $T$  sull'asse temporale. Si possono definire per il portafoglio due serie a seconda della base utilizzata, similmente alla definizione dei numeri indice a base fissa e a base mobile: sia  $n_t$  il numero di contratti in essere all'epoca  $t = 0, 1, \dots, T$ , da cui  $n_0 = n$ , e sia  $m_t$  il numero di recessi liquidati tra le epoche  $[t - 1, t[$  per  $t = 1, 2, \dots, T$ . Si definisce il *tasso di recesso a base fissa* come<sup>74</sup>:

$$\lambda_{t,0} = \frac{m_t}{n_0} ,$$

mentre si definisce il *tasso di recesso a base mobile*, sempre per  $t = 1, 2, \dots, T$ , come:

$$\lambda_t = \frac{m_t}{n_{t-1}} .$$

I due tassi differiscono appunto per la scelta della base o, meglio, degli esposti al rischio di recesso. A differenza dei numeri indice non è possibile passare dalla base fissa alla base mobile se non sono noti anche i decessi nei diversi anni, poiché il numero di assicurati dipende anche dalla mortalità del portafoglio. La seconda formula può essere modificata sostituendo al denominatore una quantità che rappresenti meglio gli esposti

---

<sup>73</sup>Il problema della definizione della formula per il tasso è sicuramente diverso per la mortalità e il riscatto: mentre per il primo il flusso da utilizzare sono solo i decessi, per il secondo, come si vedrà, possono essere presi diversi flussi inerenti al riscatto. La somiglianza sta nel fatto che, sebbene la quantità da utilizzare sia unica per la mortalità, le definizioni possono essere differenti.

<sup>74</sup>Si noti che la variabile  $t$  indica semplicemente l'antidurata, non si sta tenendo conto ancora dell'evoluzione temporale. Inoltre, le formule che verranno via via presentate utilizzano le lettere minuscole per rappresentare le diverse quantità, poiché è usuale indicare con le minuscole le quantità già osservate, determinazioni dei numeri aleatori ad essi associati; per questo si parlerà più giù di stime e non stimatori.

effettivi al rischio, dato che gli assicurati si assume possano recedere in qualsiasi istante intermedio tra le due epoche e di conseguenza la base si modifica dopo ogni istante in cui un contratto si estingue, inoltre gli assicurati sono esposti anche al rischio di decesso e nel caso di polizze con diversa durata si potrebbero avere nell'intervallo considerato uscite per termine naturale della polizza. Si può allora ridefinire la formula tramite una esposizione “centrale”  $n_{t-1}^c$ , similmente a quanto accade per il *tasso di mortalità centrale*. Nel caso di base mobile l'epoca precedente, si definisce il *tasso di recesso centrale*:

$$\lambda_t^c = \frac{m_t}{n_{t-1}^c} .$$

L'esposizione centrale è più complessa da calcolare e le differenze tra i due tassi potrebbero non essere rilevanti, qualora ad esempio i recessi osservati siano pochi e la frequenza temporale breve, per esempio settimanale. Una possibile semplificazione dell'esposizione centrale la si può ottenere supponendo ad esempio che tutti i recessi si distribuiscano uniformemente nell'anno.

Un'ulteriore modifica delle due formule si ottiene sostituendo al numero di recessi liquidati il numero di richieste di recesso effettuate, come osservato sopra. La differenza tra i due flussi è sottile ma sostanziale e il motivo risiede nel fatto che, mentre il primo dipende solamente dalla scelta dell'assicurato, quindi da una variabile esterna all'impresa, la seconda dipende anche in parte da come l'assicurazione stessa gestisce le richieste, dai tempi e dai processi interni. Le assicurazioni sono inclini a rendere più rapida possibile la procedura di liquidazione, questa però può subire dei rallentamenti in presenza di incrementi inattesi nella richiesta di riscatti. Il ritardo generato va ad impattare sul secondo flusso, la serie dei riscatti osservati, e può rendere più difficile studiare la o le cause scatenanti l'incremento. Di contro, tuttavia, il primo tipo di flusso è legato all'istante in cui si verificano gli effettivi esborsi per l'impresa – nel caso in cui ci si riferisca alle polizze con diritto di riscatto – e può pertanto facilitare lo studio dell'impatto economico dovuto ai recessi.

È pratica comune rappresentare i tassi in termini annui, anziché sulla base del periodo con cui sono stati calcolati, in modo tale da rendere confrontabili quantità calcolate con diverse frequenze; l'anno, poi, è l'unità di misura del tempo utilizzata dalla Direttiva. Fissata una certa base periodale inferiore all'anno - come la settimana - rappresentabile tramite la frequenza  $f$ , definita come il numero di volte con cui il periodo si ripete durante l'anno - per la settimana è approssimativamente 52 volte - , si definisce  $\lambda_s^{(f)}$  il *tasso di recesso con frequenza  $f$*  per  $s = 1, 2, \dots, (Tf)$ <sup>75</sup>. Per ottenere ora il tasso su base annua  $\lambda_s^{(1)}$ , cioè con frequenza pari ad uno, si applica la formula:

$$\lambda_s^{(1)} = f \lambda_s^{(f)} .$$

---

<sup>75</sup>La frequenza settimanale esatta è  $\frac{365}{7} \approx 52.14286$ , o per gli anni bisestili  $\frac{366}{7} \approx 52.28571$ , ma essendo non intere si deve approssimare al naturale più prossimo. Inoltre, la definizione potrebbe essere applicata anche per basi superiori all'anno, ma sono poco utilizzate.

Utilizzando il tasso  $\lambda_s^{(1)}$ , è possibile calcolare un tasso medio annuo, per un fissato anno  $t$ , come una media ponderata dei tassi riferiti a quel particolare anno, pesati per le rispettive esposizioni. Si definisce pertanto il *tasso di recesso medio annuo* per l'anno  $t$ -esimo, con  $t = 1, 2, \dots, T$ , come:

$$\bar{\lambda}_t = \sum_{s=1}^f \lambda_{(t-1)f+s}^{(1)} \frac{n_{(t-1)f+(s-1)}^{(f)}}{n_{t-1}},$$

con  $n_t = \sum_{s=0}^{f-1} n_{tf+s}^{(f)}$  e  $n_s^{(f)}$  il numero di contratti in essere all'epoca  $s$ . La formula può essere semplificata utilizzando le definizioni date ed ottenendo:

$$\bar{\lambda}_t = f \frac{\sum_{s=1}^f m_{(t-1)f+s}}{\sum_{s=0}^{f-1} n_{(t-1)f+s}^{(f)}}.$$

Un ulteriore tasso che può essere definito è il *tasso di recesso cumulato*, per un fissato periodo. Considerato ad esempio la durata dei contratti  $T$ , il tasso cumulato lo si ottiene come somma dei tassi di recesso a base fissa, semplificabile poi come rapporto tra il totale dei recessi verificatisi sulla numerosità totale del portafoglio:

$$\lambda_{cum} = \sum_{t=1}^T \lambda_{t,0} = \frac{\sum_{t=1}^T m_t}{n_0}.$$

Poiché il numero di recessi massimo possibile per un portafoglio chiuso è pari a  $n_0$ , il valore del tasso cumulato è compreso tra zero e uno. Sotto ipotesi di omogeneità e indipendenza degli assicurati in un *pool* chiuso - suddivisione spesso chiamata con il termine inglese "*Homogeneous Risk Group*", in breve HRG - i tassi di recesso con diversa base e il tasso cumulato posso essere visti come una stima della probabilità di recesso per le diverse antedure<sup>76</sup>. Nello specifico, è possibile vedere  $\lambda_{t,0}$  come una stima della probabilità che una polizza rimanga attiva fino a  $t-1$  e venga terminata in  $t$  per recesso dell'assicurato,  $\lambda_t$  invece come una stima della probabilità condizionata di recesso in  $t$ , sapendo che la polizza è rimasta attiva fino a  $t-1$ , e infine  $\lambda_{cum}$  come una stima della probabilità che un assicurato receda in una qualunque epoca prima della scadenza.

Le definizioni date si basano sull'ipotesi fondamentale di contratti stipulati alla stessa epoca, ovvero una generazione di contratti. Quando si intende valutare il flusso delle estinzioni ad una granularità temporale più fine, tale da rendere infattibile la suddivisione in generazioni, si può far cadere tale ipotesi e calcolare la serie a partire da un portafoglio aperto di un gruppo di contratti. Affinché l'analisi del tasso rimanga sensata, è necessario anche qui delimitare la valutazione alle HRG prese singolarmente. I tassi definibili sono sempre quello a base mobile e quello a base fissa, per un assegnata epoca di riferimento. Se il numero di entranti è approssimativamente costante nel

---

<sup>76</sup>Con "*omogeneità*" si intende che le caratteristiche peculiari dei soggetti e dei loro contratti - età, capitale assicurato, profilo di rischio, etc. - sono sufficientemente simili da poter associare alle singole probabilità di recesso la medesima distribuzione.

tempo, è possibile analizzare direttamente la serie del numero di riscatti osservati, nel caso in cui ad esempio il modello utilizzato sia un processo di *Poisson Omogeneo* - in caso contrario è però possibile passare ai modelli di *Poisson Non Omogenei*, in particolare ai processi di *Cox*, come discusso in precedenza<sup>77</sup> - .

L'interpretazione del tasso non può essere più quella di probabilità di uscita per riscatto per una certa antidurata, poiché non si ragiona più per generazioni. Quello che viene fatto è utilizzare la serie per prevedere l'evoluzione futura dei tassi e tramite questa previsione ottenere delle stime per le probabilità di riscatto nelle epoche future, per i diversi contratti del portafoglio fissato. È importante notare, inoltre, che l'utilizzo dei tassi valutati su portafogli aperti permette di introdurre nell'analisi l'evoluzione temporale della serie: le previsioni future, infatti, sono delle stime delle probabilità di riscatto per un particolare epoca futura, non per una certa antidurata, la quale può però comunque essere introdotta nei modelli per le previsioni, data la sua importanza nell'analisi, come già osservato.

Una ulteriore tasso molto utile per l'analisi dei riscatti è il cosiddetto *indice di riscatto*, riportato ad esempio nelle pubblicazioni di ANIA: “*ANIA trends flussi e riserve Vita*”<sup>78</sup>. In queste stesse viene data la sua definizione come di “*rapporto tra importi pagati per riscatto (parziale e/o totale) e l'ammontare medio delle riserve tecniche*”<sup>79</sup>. In formule è così rappresentabile: preso un portafoglio, per semplicità aperto, di contratti omogenei, valutato in una fissata epoca  $t$  e di numerosità  $n_t$ , dato  $c_t$  il totale degli importi pagati per i riscatti avvenuti tra le epoche  $[t-1, t]$ , ottenuto come somma dei valori di riscatto pagati ai recedenti  $c_t^i$ , con  $i = 1, \dots, m_t$  - quest'ultimo definito come sopra - ovverosia:  $c_t = \sum_{i=1}^{m_t} c_t^i$ , data ancora  $v_t$  la riserva tecnica totale all'epoca  $t$ , ottenuta come somma delle riserve tecniche dei singoli contratti  $v_t^i$ , per  $i = 1, 2, \dots, n_t$ , ovverosia:  $v_t = \sum_{i=1}^{n_t} v_t^i$ , si definisce l'*indice di riscatto* ad epoca  $t$  per un dato portafoglio come :

$$r_t = \frac{c_t}{v_{t-1}} .$$

La differenza fondamentale tra questo indice e i tassi sopra definiti sta nel fatto che questi ultimi si basano sul conteggio delle “*teste uscenti*”, ovvero il numero di assicurati che decidono di recedere, mentre il primo si basa sull'impatto monetario del recesso. È di fatto una serie economica, la quale perciò si presta più naturalmente ad una modellizzazione tramite le serie storiche economiche, rispetto al tasso calcolato su teste, che si presta invece a modelli di tipo sopravvivenza. Se il portafoglio sotto analisi mostra una significativa disomogeneità nei valori di riscatto, sarà più appropriato l'uso dell'indice di riscatto, poiché la dinamicità di un soggetto è correlata col valore di riscatto del suo contratto. Conseguentemente, due periodi con la stessa frequenza di riscatto potrebbero avere un diverso impatto economico quando ad esempio i tassi d'interesse per quei due periodi siano diversi.

<sup>77</sup>Per tali modelli si faccia riferimento sempre al libro: Brémaud (2020).

<sup>78</sup>Le diverse pubblicazioni sono consultabili sul sito di ANIA, alla pagina: [https://www.ania.it/pubblicazioni/-/categories/52456?p\\_r\\_p\\_resetCur=true&p\\_r\\_p\\_categoryId=52456](https://www.ania.it/pubblicazioni/-/categories/52456?p_r_p_resetCur=true&p_r_p_categoryId=52456) .

<sup>79</sup>Stefano Trionfetti (2024).

Si può ricavare una relazione tra il tasso di riscatto e l'indice di riscatto tramite l'uso delle medie campionarie dei flussi:  $\bar{c}_t = \frac{\sum_{i=1}^{m_t} c_t^i}{m_t} \iff m_t \bar{c}_t = \sum_{i=1}^{m_t} c_t^i$  ed ugualmente per  $v_{t-1}$ , questo per  $m_t$  e  $n_{t-1}$  diversi da zero. Si può allora ricavare la relazione per l'indice di riscatto:

$$r_t = \frac{c_t}{v_t} = \frac{\sum_{i=1}^{m_t} c_t^i}{\sum_{i=1}^{n_{t-1}} v_{t-1}^i} = \frac{m_t \bar{c}_t}{n_{t-1} \bar{v}_{t-1}} = \lambda_t \frac{\bar{c}_t}{\bar{v}_{t-1}} .$$

Il secondo fattore può essere approssimato osservando che il valore di riscatto è definito tipicamente come una percentuale della riserva matematica del contratto, si può perciò supporre che  $\bar{c}_t \approx \theta \bar{v}_{t-1}$ , dove  $\theta$  è l'aliquota applicata alla riserva per ottenere il valore di riscatto, compresa tra zero e uno, e assunta qui costante nel tempo e tra gli individui del *pool*. In tal modo si semplifica la frazione e si ottiene:

$$r_t \approx \lambda_t \theta \iff \lambda_t \approx r_t \frac{1}{\theta} .$$

Perciò è possibile ottenere una approssimazione della stima della probabilità di riscatto per una data epoca, noto  $\theta$ . Da questo si ricava anche che l'indice  $r_t$  è parimenti compreso tra zero e uno. Bisogna per questo fare attenzione nell'applicazione dei modelli per le serie storiche all'indice.

### 3.1.6 L'approccio tramite *Dynamic Lapse Multiplier* per l'analisi del *lapse rate*

A conclusione del paragrafo, tornando ai possibili approcci per la valutazione del *lapse rate*, si vuole riportare un ultimo metodo che viene correntemente utilizzato, in particolare nel settore americano, basato sul cosiddetto “*Dynamic Lapse Multiplier*”, anche detto “*Dynamic Adjustment Factor*”<sup>80 81</sup>. Come ben riassunto nell'articolo Fabio Bazione (2019), i modelli per il tasso di estinzione anticipata possono essere suddivisi in modelli “*ottimali*” e “*sub-ottimali*”, a seconda che si ipotizzi che gli agenti siano perfettamente razionali e recedano nell'istante ottimale o che siano imperfetti nella loro scelta. La seconda tipologia di modelli può essere ulteriormente suddivisa in modelli “*one-stage*” e in modelli “*two-stage*”: i primi sono costruiti tenendo assieme la componente statica e quella dinamica, e tra questi fanno parte tutti i modelli sopra elencati; i secondi invece separano l'analisi delle due componenti, statica e dinamica, in particolare viene definita una funzione, che può essere deterministica, per rappresentare

---

<sup>80</sup>Un altro metodo che potrebbe venir presentato è quello basato sulle tecniche di *machine learning*, queste però, per la presente Tesi, non sono d'interesse, data la loro natura di *black box*, e quindi per le quali non è possibile giustificare le stime ottenute o valutare il legame con le variabili esogene. Su una applicazione del *machine learning* per la modellizzazione del tasso di riscatto si rimanda all'articolo Azzone et al. (2022).

<sup>81</sup>Con riferimento all'utilizzo nel settore assicurativo americano di tale approccio, si veda come esempio il documento: *Considerations Regarding Dynamic Lapses in Actuarial Modeling* (2023).



la componente di base, o idiosincratice, la quale è poi moltiplicata per un fattore dipendente dal tempo e da variabili macroeconomiche, detto appunto *Dynamic Lapse Multipliers*. La prima componente è dipendente da quelle che sono le variabili particolari del contratto, come l'età e l'antidurata, e rappresenta quello che è il comportamento base o statico, rappresentabile tramite una funzione deterministica o stocastica, come una regressione non lineare; inoltre, il fattore si riferisce ad un singolo assicurato. La seconda componente, che dà il nome al metodo, è anch'essa una funzione, la quale però dipende dalle variabili macroeconomiche, in particolar modo dalla differenza tra il tasso di rendimento di mercato e il tasso accreditato al contratto, ed è quindi stocastica; a differenza della prima componente viene presa uguale per ogni assicurato. Questo fattore è classicamente definito tramite una curva ad esse, la "*S-Shaped Curve*", così come chiamata in letteratura, ovvero una funzione monotona non decrescente rispetto alla differenza tasso di mercato-tasso accreditato, che assume valori tra zero ed un massimo - inferiore all'inverso del primo fattore, per garantire che il tasso sia inferiore ad uno - . Un documento che riporta alcune possibili forme funzionali è il Baione, Biancalana, and Angelis (2021) - oltre a presentarne una sua propria - , tra queste si trova la curva logistica, ma sono possibili anche funzioni discontinue, come quelle a gradini<sup>82</sup>.

La forza di questo tipo di procedure è quella di poter suddividere l'analisi e quindi gestire separatamente la valutazione della parte statica e della parte dinamica, oltre ad ottenere un modello molto flessibile grazie alla libertà di scelta delle funzioni e per il quale l'interpretazione è più intuitiva. La scelta delle curve però è limitata a quelle in una sola variabile, tipicamente la differenza tasso di mercato-tasso accreditato, questo perché rende più semplice la costruzione del modello; inoltre, come visto nel capitolo precedente, alcune variabili contrattuali sono correlate a comportamenti più dinamici, mentre qui comparirebbero come elementi separati.

L'insieme delle procedure appena viste costituisce un sunto dell'attuale corpus di strumenti utilizzati per la modellizzazione del *lapse rate*. Essi, tuttavia, non si presentano adatti alla valutazione del *mass lapse* e le ragioni di ciò verranno argomentate nel prossimo paragrafo.

### 3.2 Il problema del *mass lapse* e le proposte in letteratura

Il *mass lapse* si caratterizza, innanzitutto, come un evento estremo o catastrofe, appartenente alle code della distribuzione del *lapse rate*. I modelli precedentemente descritti permettono di ottenere una stima della distribuzione del tasso. Tramite questa si potrebbe, quindi, procedere o stimando l'impatto dell'evento di massa, in termini di

---

<sup>82</sup>Si faccia riferimento sempre al Baione, Biancalana, and Angelis (2021) per la definizione della "*S-Shaped Curve*", ma si può trovarla discussa anche negli articoli Barsotti, Milhaud, and Salhi (2016) e Loisel and Milhaud (2011).



numero di abbandoni, utilizzando ad esempio il Valore a Rischio della distribuzione stimata, per il fissato livello di confidenza del 99.5% della Direttiva, oppure stimando la probabilità di un aumento dei riscatti del 40%, come da formula standard<sup>83</sup>. Tuttavia il ragionamento è fortemente problematico: tali modelli non sono mirati all'analisi degli eventi estremi, l'intento è sempre quello di ottenere delle buone stime dei valori attesi o di un intervallo di valori e perciò tendono spesso a sottostimare le valutazioni delle code della distribuzione. Questo accade ad esempio nei modelli parametrici dove si ipotizzano errori a distribuzione *Normale*, la quale è caratterizzata da code leggere. Passando pure ai modelli non parametrici, i quali prevedono la stima diretta della FdR, si osserva, anche in campioni di grandi dimensioni, che le code risultano sottostimate. Questo avviene perché gli eventi estremi si possono manifestare con una frequenza troppo bassa per poter raggiungere stime accettabili della loro probabilità<sup>84</sup>. A questo problema se ne aggiunge poi un secondo: è usuale trattare gli eventi estremi come *outliers*, cioè come dati che “cadono fuori” dall'andamento medio generale di un campione. Di fronte a questi valori estremi, si è molto spesso portati alla loro rimozione dall'analisi o ad un loro trattamento separato, con l'intento di migliorare il *fitting* del modello. Il risultato è un modello che lavora bene in generale, ma non nella situazione particolare del *mass lapse*.

Per far fronte a questo tipo di problematiche si possono prendere più strade, due in particolare: la prima è di modificare il modello affinché possa meglio rappresentare quelli che sono gli eventi estremi, ad esempio utilizzando una diversa distribuzione degli errori, a code più pesanti, o analizzando a fondo la natura dell'evento catastrofale e provando a riprodurla in termini matematici; la seconda, applicabile a tutti i modelli - sotto condizione che valgano particolari ipotesi - è di far uso della “*Teoria dei Valori Estremi*” (“*Extreme Value Theory*”, EVT<sup>85</sup>). Tale teoria rientra nella famiglia dei modelli semiparametrici e si basa su una proprietà asintotica dei valori estremi, tramite la quale è possibile ottenerne una distribuzione limite per essi. Questo tipo di approccio è largamente usato per studiare fenomeni simili al *mass lapse*, ciò che rende però questo particolare caso più complesso è il fatto che la teoria base richiede due ipotesi fondamentali: l'indipendenza e l'identica distribuzione dei numeri aleatori associati alle osservazioni del campione. Per la serie dei tassi di riscatto le due ipotesi possono essere problematiche. Partendo dall'identica distribuzione, caso meno critico tra i due, questa non può essere assunta perché il tasso di riscatto dipende significativamente dalle variabili discusse precedentemente. Ad esempio, una variazione dei tassi, potrebbe implicare una variazione del valore atteso della distribuzione dei riscatti. Spostandosi invece sull'ipotesi di indipendenza, la questione si fa più delicata. Vi sono infatti due situazioni particolari, in parte già discusse, in cui si può manifestare un fenomeno di

---

<sup>83</sup>Parlando di distribuzione del tasso di riscatto si sta semplificando la questione, si dovrebbe più propriamente fare un discorso in termini temporali.

<sup>84</sup>detto in altri termini, la coda della distribuzione è sotto campionata e di conseguenza lo stimatore rimane lontano dal risultato asintotico.

<sup>85</sup>Per una introduzione alla Teoria dei Valori Estremi, si può fare riferimento al libro: Coles et al. (2001).

correlazione<sup>86</sup>. La prima situazione è riferita alla modellizzazione del *lapse rate* tramite processi ARMA, in questi infatti è insita la natura di dipendenza dai ritardi del processo stesso - l'*autocorrelazione* - ; mentre, invece, la seconda situazione è più specifica del fenomeno del *mass lapse*, poiché legata a quell'effetto, esposto nel precedente capitolo, di contagio, e quindi di correlazione, tra gli individui, che si verifica durante periodi particolari come quelli di crisi economica. Sono in realtà proposte delle generalizzazioni della teoria che permettono di utilizzare dati non *iid*, ad esempio utilizzando serie stazionarie - la stazionarietà è però problematica per la serie dei riscatti, come verrà detto poco sotto - , e per dati dipendenti da variabili esterne. Il più volte citato Biagini et al. (2021) fa utilizzo di questi modelli proprio per l'analisi del rischio di estinzione di massa.

Tornando al problema della correlazione tra i soggetti in periodi di crisi, esso è anche il motivo per il quale molti degli altri modelli visti in precedenza per il tasso di riscatto non sarebbero adatti per determinare l'impatto di un evento di massa. Prendendo i GLM come esempio, i quali costituiscono la procedura più utilizzata tra le elencate, essi richiedono, come prima ipotesi, che sia verificata l'indipendenza delle unità statistiche.

Questo problema dell'indipendenza è, come già accennato, risolvibile - almeno in parte - ricorrendo ai modelli ARMA, ma l'utilizzo di questi fa sorgere un'ulteriore problematica: la "Non Stazionarietà". Un'ipotesi fondamentale, infatti, richiesta da questi modelli è la "Stazionarietà del Processo", con cui si intende che la serie debba mantenere un comportamento simile nel corso del tempo, in modo tale che stimare la serie in un punto permetta di conoscere il suo comportamento passato e futuro. Essa non può essere assunta nel caso del tasso di recesso, proprio perché il fenomeno del *mass lapse* si manifesta come una rottura in un certo istante di tempo dell'andamento medio del tasso; tale frattura causa appunto la caduta dell'ipotesi e quindi la non stazionarietà. Perciò, l'utilizzo delle serie storiche per l'analisi del tasso di riscatto richiede di poter inserire la dinamica del salto causato dall'evento di massa in modo tale che il processo risultante sia stazionario<sup>87</sup>.

Un'ultimo problema che si può citare è la scarsità di osservazioni riguardanti il fenomeno, proprio per la sua natura di evento estremo. Non poter disporre di campioni sufficienti complica di molto la sua stima: *come si può stimare un fenomeno senza poterlo osservare?* È necessario introdurre nuove ipotesi sulla natura dell'evento e sulla dinamica che provoca sulla serie, le quali potrebbero essere difficilmente giustificabili se non tramite l'*expert judgement*.

Ricapitolando, il rischio di estinzione di massa presenta queste quattro grandi criticità: l'assenza di indipendenza e di identica distribuzione, la natura di evento estremo, il causare la non stazionarietà della serie dei riscatti e l'assenza di dati in merito. An-

---

<sup>86</sup>Si sta passando qui dal termine "*dipendenza*" al termine "*correlazione*", è importante ricordare che dal punto di vista matematico i due non sono sinonimi: l'indipendenza implica l'incorrelazione e non il contrario, ma la correlazione implica la dipendenza, per questo è possibile il passaggio fatto nel discorso

<sup>87</sup>Per la definizione di "Stazionarietà" e per le problematiche causate della non stazionarietà si rimanda alle Appendici.

che la letteratura scientifica in merito attualmente è molto scarna, si può fare menzione però di quattro articoli che trattano nello specifico dell'argomento. Si intende ora presentarli di seguito, sottolineando in particolare le soluzioni che sono state elaborate per far fronte alle criticità osservate.

### 3.2.1 Biagini, Huber, Jaspersen e Mazzon 2020

Il documento affronta il problema della modellizzazione del rischio di recesso in massa rifacendosi alla “Teoria dei Valori Estremi”, in tal senso: poiché non è possibile assumere che il processo dei tassi di abbandono sia *iid* viene applicata una variante di questo, chiamata “*Dynamic Peaks Over Threshold Method*” (DPOT). L'articolo applica tale metodo a due dataset di tipo panel, riferiti a due mercati distinti, americano e tedesco, contenenti osservazioni per diverse imprese, nel tempo. Il primo panel, riguardante il mercato americano, ha 617 compagnie, osservate dal 1996 fino al 2018, mentre il secondo, riferito al mercato tedesco, ha 122 compagnie, osservate dal 1996 al 2018. Tali campioni sono stati ritenuti correlati sia nel tempo che nello spazio a causa di effetti dovuti a variabili o fattori dipendenti dal tempo o dalla compagnia, come ad esempio l'anno di osservazione e la dimensione del portafoglio. Il DPOT permette di introdurre queste come covariate, oltre ad altre possibili variabili di controllo, per ottenere la convergenza alla stessa distribuzione del POT senza però necessitare dell'ipotesi di indipendenza e identica distribuzione. I risultati ottenuti dall'applicazione del modello ai campioni permettono di asserire, come già era stato detto nel capitolo precedente, che la valutazione fatta da *Solvency* per il *mass lapse risk* risulti inadeguata, poiché il rischio è fortemente dipendente dai diversi fattori succitati. Risulta, ad esempio, dal campione tedesco, che la distribuzione stimata delle frequenze di riscatto riporti come 99,5-esimo percentile solo un valore del 25%, mentre quello americano riesce a superare il 40%.

### 3.2.2 Loisel e Milhaud 2011

L'articolo non ha come tema di riferimento il *mass lapse* esplicitamente, ma si propone di definire un modello stocastico a sostituzione della curva deterministica utilizzata nell'approccio, già discusso, del *Dynamic Lapse Multipliers*.

Il modello definito tiene conto anche del rischio di estinzione di massa, perché presenta al suo interno il meccanismo di contagio degli assicurati nella scelta di riscattare o meno. Permette, inoltre, al fenomeno del contagio di evolvere nel tempo per effetto di variabili esterne, come il tasso di mercato.

Le ipotesi alla base del ragionamento sono le seguenti: il comportamento dell'assicurato, in condizioni di mercato “stabile”, si presenta come non correlato e con tasso di abbandono medio positivo relativamente piccolo e la cui distribuzione è associabile ad una *Gaussiana*. In caso di mercato “non stabile” subentra il meccanismo precedentemente citato del *Copycat Behavior*: la scelta di recedere dal contratto o no dipende con una certa probabilità dalle scelte degli altri assicurati, e più aumenta l'instabilità più

umenta la probabilità tale probabilità. La distribuzione del tasso di abbandono che ne risulta è una cosiddetta distribuzione “bimodale” - il modello implicato sottostante è spesso detto “*0-1 Model*” - , nel quale cioè vi sono due mode, ad indicare che o gli assicurati recederanno quasi in massa o non recederà quasi nessuno.

L’articolo passa poi ad un’applicazione del modello usando parametri realistici, ricavati da un portafoglio di una compagnia di riassicurazione spagnola, osservata mensilmente da febbraio 2002 a dicembre 2007, con dimensioni che variano da 291 contratti iniziali a 25766 finali, in più è imposta un’ipotesi di massima percentuale di riscatti possibile mensile del 3,5%. I risultati ottenuti verificano l’ipotesi di distribuzione Gaussiana del *lapse rate* nella condizione in cui il mercato si trovi relativamente stabile. Poiché poi il dataset non contiene osservazioni attribuibili a mesi con tassi di abbandono estremi, gli autori riportano semplicemente il VaR e TVaR stimati per diversi gradi possibili di correlazione e di propensione all’abbandono rispetto allo *spread* tra tasso di mercato e tasso accreditato.

### 3.2.3 Barsotti, Milhaud e Salhi 2016

L’intento degli autori del documento è quello di coniugare il fenomeno dell’effetto di contagio con il comportamento dinamico. Questa unione ha come idea quella di poter modellizzare l’andamento dinamico del *lapse rate* tenendo conto dei possibili eventi di estinzione di massa.

Nello specifico, l’ipotesi matematica alla base è di supporre che l’intensità di abbandono al tempo  $t$ , parametro associato al processo di conta che descrive il numero di abbandoni fino ad un dato istante, sia a sua volta un processo stocastico, chiamato “a contagio dinamico”; tale processo è una modifica dei processi cosiddetti di *Hakwes*. Detto tout court, il processo cattura il possibile - aleatorio - passaggio da un’intensità di abbandono “strutturale”, stazionaria nel tempo, dovuta a fenomeni statici, ad una intensità “dinamica”, dovuta invece alla somma di due effetti distinti: quello dovuto alle scelte razionali dell’assicurato, che dipenderanno dall’evoluzione dei tassi d’interesse, e quello dovuto piuttosto al contagio tra gli assicurati.

La parte dipendente dai tassi di interesse è costruita su un processo di conta, a partire da un “Moto Browniano Geometrico” per l’evoluzione del tasso forward istantaneo più l’evoluzione - deterministica - del tasso a credito; quando la differenza percentuale tra i due supera una certa soglia, il processo di conta ha un salto, di ampiezza aleatoria. Il secondo pezzo è anch’esso un processo di conta, che fa dipendere l’intensità dai recessi che si sono verificati in precedenza.

L’ultima parte dell’articolo è dedicata ad un confronto numerico tra la standard formula di *Solvency*, l’approccio tramite moltiplicatore dinamico e il modello definito dagli autori. I dati sono ottenuti in via simulativa, senza l’utilizzo di un campione reale, ed i parametri sono fatti variare per ottenere diversi scenari per diversi confronti.

### 3.2.4 Cheng, Hilpert, Lavasani e Schaefer 2023

Come nell'articolo precedente, anche qui alla base si può trovare una modifica dei processi di *Hakwes*, con l'intento di riprodurre l'effetto di correlazione tra gli assicurati. In aggiunta, il documento va a misurare l'impatto dell'evento estremo sull'*Equity* e determina gli adeguamenti necessari per tener conto nel *pricing* del prodotto di tale rischio. Dal punto di vista matematico, viene definita un'impresa ideale che opera su un mercato privo di arbitraggi, con un portafoglio di polizze contratte da assicurati suddivisibili in due sottogruppi: i "professionali", il cui comportamento è quello di un "agente razionale" e i "non professionali", che agiscono sotto altri criteri personali; i primi abbandonano il contratto quando risulta "ottimale" l'opzione di riscatto, i secondi invece in base a scelte personali, che potrebbero anche dipendere dalle scelte degli altri assicurati, per effetto del contagio che esiste tra questi. Anche qui il processo "*numero di abbandoni nel tempo*" dipende da una parte dinamica che segue le condizioni di mercato, da una parte stabile nel tempo e da un effetto di contagio e correlazione seriale; il modello stocastico così definito è una variazione di quello presentato nel Barsotti, Milhaud, and Salhi (2016).

Viene poi descritto il comportamento del processo degli attivi tramite un processo stocastico con "salti", dovuti al processo di conta del numero di abbandoni e di ampiezza pari ad un processo ulteriore che rappresenta l'evoluzione del beneficio pagato a riscatto, che a sua volta dipende dalle garanzie contrattuali e dal tasso di rivalutazione. Come ultima componente aggiunge la valutazione del prezzo dell'opzione di riscatto, all'interno della cornice matematica definita.

La parte conclusiva svolge un'analisi simulativa del modello. Alcuni parametri sono stati calibrati su un tipo di polizze vendute attualmente dalla compagnia canadese *Great West Life*, in modo tale da ottenere quantità realistiche. I restanti parametri sono invece valori calibrati per essere più o meno realistici. L'approccio utilizza il "*Least Squares Monte Carlo Method*", non disponendo di soluzioni in forma chiusa.

Questi sono quattro approcci attualmente riportati in letteratura. La presente Tesi si proporrà, nel seguente capitolo, di fornire un ulteriore possibile approccio alla modellizzazione del rischio.

## 4 I modelli *Regime Switching*

I modelli proposti dalla presente ricerca per l'analisi del *mass lapse* appartengono ad una famiglia di modelli parametrici per le serie storiche, chiamata "*Regime Switching Model*" (RSM). L'idea alla base che accomuna tutti questi modelli è quella di avere un processo stocastico (*processo dipendente*), solitamente un processo Autoregressivo di ordine  $p$ , i cui parametri sono processi stocastici a loro volta e le cui determinazioni dipendono da un processo stocastico (*processo indipendente*), che può essere noto o non noto, e endogeno o esogeno al modello<sup>88</sup>. In particolare, i parametri prendono valori in un insieme discreto e ogni determinazione del vettore dei parametri rappresenta un particolare stato o "regime" del processo. Ad esempio quando il processo indipendente supera una data soglia, il processo dipendente transita ad un'altro regime, da cui il nome della famiglia. Gli RSM fanno parte della classe dei modelli non lineari, si può verificare infatti che i processi risultanti non hanno *trend* lineari, ma al massimo lineari a tratti. Risultano quindi processi non stazionari, la cui stima dei parametri però non soffre dei problemi causati dalla non stazionarietà.

Questa classe di modelli sono ampiamente studiati in letteratura ed applicati, anche al di fuori della materia finanziaria, sebbene però siano ancora poco implementati nei *software* classici, come R<sup>89</sup>. Il loro ambito di utilizzo classico è quello dell'analisi di fenomeni dinamici, cioè processi la cui struttura varia stocasticamente nel tempo, un effetto solitamente chiamato con il termine inglese di "*time varying*". La differenza tra i regimi può essere più o meno pronunciata e si adatta bene a modellizzare serie che osservano periodi con valori molto distinti, come accade per i tassi di rendimento di mercato o per l'inflazione. Sono utilizzati anche per studiare fenomeni che presentano particolari forme di ciclicità, nell'idea appunto che il passaggio tra regimi si manifesti con una certa frequenza dipendente da un fenomeno sottostante. Tali modelli presuppongono nello specifico che esista una relazione non diretta tra processo dipendente e processo indipendente, ma solo in funzione di un particolare evento soglia relativo al secondo processo.

Come detto, in generale la famiglia fa parte di una macrofamiglia di modelli detti "non lineari", per i quali la relazione funzionale che gli definisce non è di tipo lineare, come accade invece per i modelli Autoregressivi. Per una discussione approfondita sul tema dei modelli non lineari per le serie storiche e nello specifico sugli RSM, si riportano Tong (1990) e Franses and Van Dijk (2000), mentre per una presentazione più mirata ma generale sulla famiglia degli RSM si consigliano Krolzig (2002) e Dijk, Teräsvirta, and Franses (2002).

---

<sup>88</sup>L'utilizzo dei termini "processo dipendente" e "processo indipendente" non è proprio della letteratura sull'argomento, ma è utilizzato qui per poter rendere più chiaro e scorrevole il discorso. La scelta dei vocaboli "in - dipendente" si rifà semplicemente alla regressione lineare, dove si parla di variabile dipendente (o "risposta") e variabili indipendenti (o "regressori").

<sup>89</sup>A questo proposito si può leggere l'introduzione di Liu, Chen, and Tsay (2020).

Il rischio di estinzione di massa può essere rappresentato tramite gli RSM, prendendo come variabile di riferimento il tasso di recesso descritto nel precedente capitolo. Il tasso si mantiene stabili su una certa struttura, un primo regime, identificabile come il “regime dei tassi bassi”. A causa di una variabile, esterna o interna, nota o non nota, si verifica una *shock* identificabile con l’evento *mass lapse*, che porta i tassi ad un secondo regime con un valore medio più alto rispetto al primo, identificabile come il “regime dei tassi alti”.

Grazie agli RSM è quindi possibile tener conto della non stazionarietà e non linearità della serie dei riscatti. I salti tra regimi catturano l’idea del *break strutturale* causato dal *mass lapse* e possono essere fatti dipendere da variabili come l’effetto di contagio, qualora si usasse come processo indipendente la serie dei riscatti stessi ma ritardata, o come la variazione dei tassi d’interesse o, ancora, farli dipendere da un qualche serie correlata ad eventi particolarmente significativi del settore assicurativo. In ottica modello interno, l’uso degli RSM permette di separare la componente catastrofale da quella non catastrofale, e permette di stimare una probabilità dell’evento *mass lapse*. Bisogna aggiungere, tuttavia, che sorgono delle problematiche nell’applicazione di questi modelli alla serie dei riscatti. La loro esposizione è rimandata però a fine capitolo.

Tra gli RSM sono stati selezionati, per essere riportati, solo alcuni, quelli in particolare ritenuti più vicini al problema e più semplici nella loro applicazione statistica. Questi sono:

- i modelli autoregressivi a *barriera* o *Threshold Autoregressive Model* (TAR),
- i modelli autoregressivi a *autoeccitanti* o *Self Exciting Threshold Autoregressive model* (SETAR),
- i modelli autoregressivi a *transizione graduale* o *Smooth Transition Autoregressive Model* (STAR).

In aggiunta, verrà proposto, come metro di paragone, un semplice modello autoregressivo con un *trend* deterministico non lineare, che riproduca il comportamento atteso della serie quando si verifica l’evento *mass lapse*. La forma funzionale del *trend* è stata scelta in base alle conoscenze ex-ante di come dovrebbe comportarsi la serie, ovvero una curva crescente fino ad un massimo e poi decrescente fino a tornare al livello consueto. Il modello in questione ha delle forti critiche, che verranno affrontate al momento della sua definizione, è però utile considerarlo perché rappresenta una *baseline* per gli altri modelli.

Tra gli RSM vi è un’importante classe di modelli, anch’essi molto studiati e utilizzati, detti i *Markov Switching Model*. A differenza dei modelli sopraccennati, i *Markov Switching Model* hanno come variabile indipendente, che determina il passaggio tra regimi, un processo non osservabile, il quale si suppone sia *Markoviano*, ovvero che dipende dalla sola determinazione passata più prossima. Si è scelto di non presentare questi modelli per due motivi: il primo, più teorico, è che, seguendo quanto appreso dalla letteratura, i passaggi tra regimi sono condizionati da variabili come il tasso d’interesse o la disoccupazione, che sono osservabili, il secondo motivo, più pratico, riguarda il fatto

che il *dataset* che sarà utilizzato per l'analisi empirica non è stimabile tramite questo modello; infatti i dati presentano un solo picco, osservato non integralmente, di contro i *Markov* RSM richiedono di un campione che presenti più *shock* possibili, osservati nella loro interezza.

Gli RSM si basano sui modelli AR, così come il processo con *trend* deterministico che si andrà a presentare si basa su questi modelli. È possibile definire delle forme equivalenti tramite i MA o gli ARIMA, modelli molto comuni tra le serie storiche, come ricordato in precedenza. Tuttavia, in generale, ma anche qui nello specifico, si preferiscono gli AR, poiché sono molto più semplici da stimare, si adattano bene al problema in esame, data la loro componente autoregressiva interpretabile come il contagio tra gli assicurati, e soprattutto, si può dimostrare, che i MA e gli AR sono interconnessi. Nello specifico, si dimostra che un processo AR stazionario è equivalente ad un processo MA di ordine infinito<sup>90</sup>.

## 4.1 I Modelli TAR

I primi modelli affrontati sono i modelli Autoregressivi a Barriera, dall'inglese *Threshold Autoregressive Model* (TAR). Per comprenderli a pieno si parte introducendo quelli che sono i modelli per i *break* strutturali permanenti di processi Autoregressivi, ovvero quei modelli che presuppongono di avere una serie temporale che segua un AR fino ad un certo istante noto  $\tau$ , dopo il quale il processo si modifica e prende la forma di un diverso AR<sup>91</sup>. Per semplicità si considerano due processi AR di ordine 1, ma il discorso è facilmente generalizzabile; il modello che si ottiene è rappresentabile in questo modo:

$$Y_t = \begin{cases} \alpha_0 + \alpha_1 Y_{t-1} + \epsilon_t & \text{per } t < \tau \\ \beta_0 + \beta_1 Y_{t-1} + \epsilon_t & \text{per } t \geq \tau \end{cases} .$$

Il primo AR ha come vettore dei parametri  $\underline{\alpha} = (\alpha_0, \alpha_1)$ , mentre il secondo ha come vettore  $\underline{\beta} = (\beta_0, \beta_1)$ ; la varianza dell'errore  $\epsilon_t$  è presa uguale per i due modelli. Se i due vettori differiscono per una qualche componente, i due modelli risultano distinti e costituiscono di fatto due regimi. Il modello può essere riscritto più comodamente utilizzando una funzione indicatrice, definita come:

$$|t \geq \tau| = \begin{cases} 1 & \text{per } t \geq \tau \\ 0 & \text{per } t < \tau \end{cases} ,$$

---

<sup>90</sup>Per una definizione di modello AR e ARIMA si rimanda alle Appendici

<sup>91</sup>Questo tipo di modelli non sono applicati ai soli processi AR.



da cui si può facilmente ricavare:

$$Y_t = (\alpha_0 + \alpha_1 Y_{t-1})(1 - |t \geq \tau|) + (\beta_0 + \beta_1 Y_{t-1})|t \geq \tau| + \epsilon_t .$$

Andando a svolgere i prodotti e facendo una sostituzione di parametri, il modello può essere anche riscritto come:

$$Y_t = a_0 + b_0 |t \geq \tau| + a_1 Y_{t-1} + b_1 |t \geq \tau| Y_{t-1} + \epsilon_t ,$$

nel quale si riconosce un modello simile ai modelli di regressione semplice con variabile *dummy*, in italiano appunto variabile indicatrice, per intercetta e con interazione.

Come ulteriore riscrittura, si possono esprimere i parametri in funzione della variabile indicatrice, a dimostrazione del fatto che questi modelli sono particolari forme di modelli per parametri *time varying*:

$$Y_t = \phi_0(|t \geq \tau|) + \phi_1(|t \geq \tau|)Y_{t-1} + \epsilon_t$$

Questo tipo di modelli è facilmente stimabile poiché basta separare il campione nei valori precedenti e successivi a  $\tau$ , avendo  $\tau$  noto, e poi si stimano separatamente i due AR riferiti ai due campioni. Il caso però più reale ed interessante è quello che prevede che l'istante di transizione sia non noto e che si voglia stimare anche quello a partire dal campione. Le opzioni sono principalmente due: la prima prevede di utilizzare un test statistico che possa individuare i possibili *break* a partire dal campione, il secondo è proprio dei modelli TAR indagati.

Partendo dal primo approccio, vengono solitamente utilizzati due particolari test per valutare le rotture di un modello: un test per *break* esogeno e un test per *break* endogeno, dove con esogeno ci si riferisce ai test che valutano, sotto un certo criterio, se una data fornita a priori è una possibile data di *break*, mentre i modelli endogeni valutano, usando un certo criterio, tutte le date del campione. Il primo, appartenente ai test esogeni, è il test di *Chow*, il quale è impostato come un normale test d'ipotesi, con ipotesi nulla l'uguaglianza dei parametri dei due modelli che si ottengono suddividendo il campione rispetto alla data fornita. Il secondo, all'opposto endogeno, è il "*Quandt Likelihood Ratio statistic*" (QLR) e si basa sul primo: viene fissato un intervallo di valori, compresi strettamente tra la prima e l'ultima data del campione; per ogni valore dell'intervallo viene fatto un *Chow* test, o meglio viene calcolata la statistica *F* data da quest'ultimo; la data a cui è associato il valore massimo, se la statistica supera una certa soglia critica fissata, è valutata come l'istante di *break*. Nel caso in cui non vi siano valori che superano la soglia fissata allora si può giudicare il campione privo di *break*. Entrambi i test hanno dei punti deboli: il primo richiede di fornire esternamente una data, mentre il secondo dipende molto dalla scelta dell'intervallo di date considerato e non è in grado di valutare la presenza di più istanti di rottura<sup>92</sup>.

---

<sup>92</sup>Sui due test appena citati si può fare riferimento a Stock and Watson (2020).

Per quanto concerne invece il secondo metodo di stima di  $\tau$ , se ne discuterà dopo l'introduzione dei modelli TAR. Ciò che li definisce è una generalizzazione del modello precedente, infatti suppone che la rottura del processo, che si può ora considerare non più una rottura ma una “transizione tra regimi”, avviene per effetto del superamento di una barriera da parte di un processo esogeno, o indipendente:

**Definizione (Modello Autoregressivo a Barriera a due regimi)** Dato un processo stocastico, esogeno,  $\{Z_t; t \in \mathbb{Z}\}$  e un numero certo  $B \in \mathbb{R}$  detto “barriera”, si definisce il processo Autoregressivo a Barriera, di ordine 1 e a due regimi ( $TAR(1, 2)$ ), il processo stocastico  $\{Y_t; t \in \mathbb{Z}\}$  tale per cui:

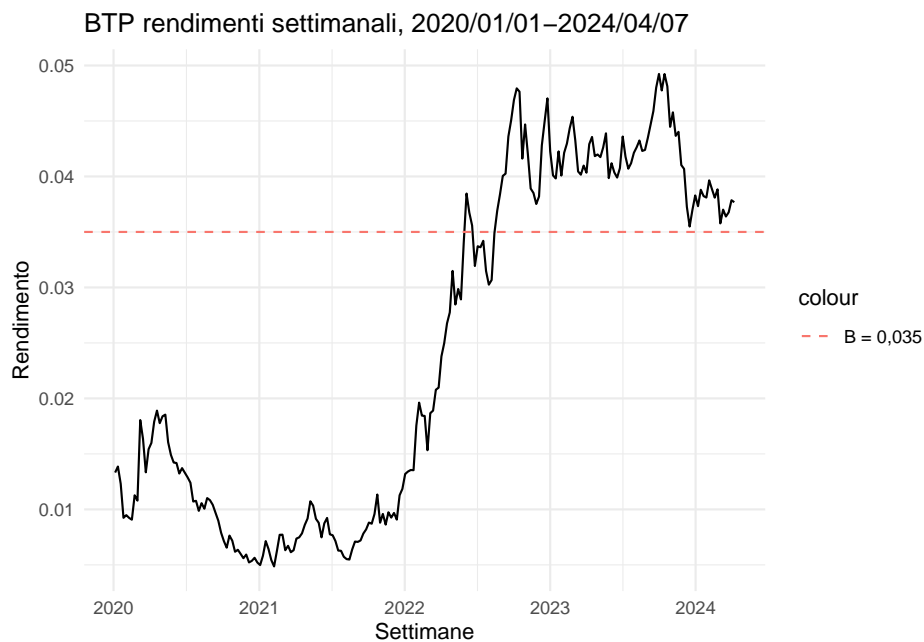
$$Y_t = \begin{cases} \alpha_0 + \alpha_1 Y_{t-1} + \epsilon_t & \text{se } Z_t < B \\ \beta_0 + \beta_1 Y_{t-1} + \epsilon_t & \text{se } Z_t \geq B \end{cases},$$

con  $\epsilon_t \sim WN(0; \sigma^2)$ , indipendente dalla storia passati di  $Y_t$ .

Anche in questo caso si può riscrivere il modello in forma compatta:

$$Y_t = (\alpha_0 + \alpha_1 Y_{t-1})(1 - |Z_t \geq B|) + (\beta_0 + \beta_1 Y_{t-1})|Z_t \geq B| + \epsilon_t$$

La funzione indicatrice in questo caso è un numero aleatorio poiché dipende dalle determinazioni del processo indipendente. Di seguito viene presentato una simulazione di un modello  $TAR(1, 2)$  con parametri  $\underline{\alpha} = (0; 0, 3)$ ,  $\underline{\beta} = (0, 04; 0, 7)$  e  $\sigma_\epsilon^2 = 0, 05$ , con processo esterno il valore dei rendimenti settimanali dei BTP italiani, dal 2020/01/01 al 2024/04/07, e barriera  $B = 0.04$ :

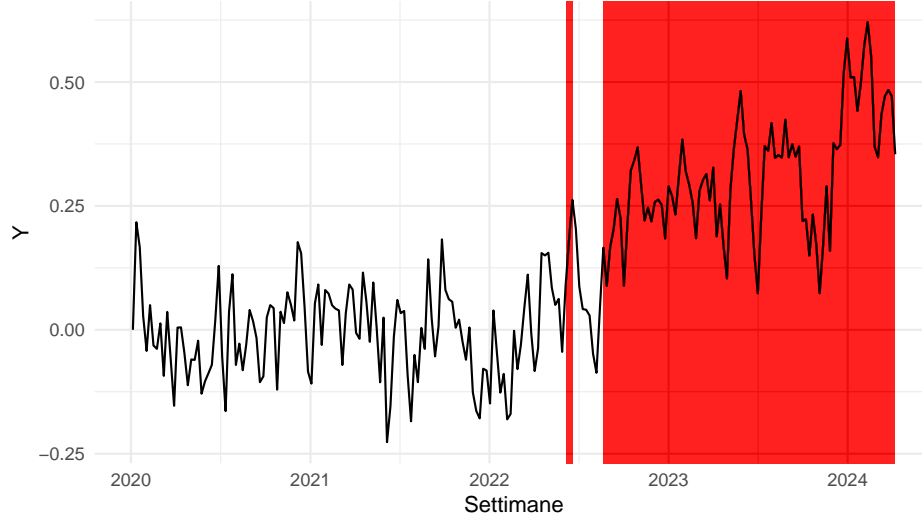


### Simulazione di un processo TAR(1,2)

Parametri:  $a=(0;0.4), b=(0,1;0,7), s_2 = 0,005$  e  $B=0,035$ .

Variabile indipendente: BTP rendimenti settimanali, 2020/01/01–2024/04/07.

Le zone in rosso rappresentano le regioni con il secondo regime.



Il secondo grafico mostra perfettamente il cambio che subisce la serie. Una cosa che si può notare è che le transizioni sono sì immediate ma non tanto quanti si potrebbero osservare in un modello di regressione semplice. Infatti la parte autoregressiva fa dipendere la serie dai valori precedenti, la qual cosa permette di decelerare il salto.

Un'aggiunta al modello precedente può essere quella di permettere ai due regimi di avere un diverso parametro per la varianza. Sempre mantenendo un unico processo degli errori, il TAR(1,2) con diversa varianza può essere scritto in questo modo:

$$Y_t = (\alpha_0 + \alpha_1 Y_t + \sigma_\alpha \epsilon_t)(1 - |Z_t \geq B|) + (\beta_0 + \beta_1 Y_t + \sigma_\beta \epsilon_t)|Z_t \geq B| ,$$

dove qui si ha  $\epsilon_t \sim WN(0, 1)$ . Si ottiene infatti per il primo regime  $\sigma_\alpha \epsilon_t \sim WN(0, \sigma_\alpha)$ , e per il secondo l'analogo con il parametro sostituito.

Il modello TAR(1,2) può essere definito in generale per un ordine  $p$  e un numero di regimi  $r$ . Ad ogni regime è associata una certa barriera  $B$ , ma il processo indipendente rimane sempre uno solo. La definizione che si dà è la seguente:

**Definizione (Modello Autoregressivo a Barriera multiregime)** Dato il processo esogeno  $\{Z_t; t \in \mathbb{Z}\}$  e dati  $B_1 < \dots < B_{r-1}$ , con  $B_i \in \mathbb{R}, i = 1 \dots r - 1$ , il processo  $\{Y_t; t \in \mathbb{Z}\}$  segue un "Modello Autoregressivo a Barriera multiregime", di ordine  $p \in \mathbb{N}$  e con numero di regimi  $r$ , se e solo se vale:

$$Y_t = \begin{cases} \phi_{0,1} + \sum_{h=1}^p \phi_{h,1} Y_{t-h} + \epsilon_t & \text{per } Z_t < B_1 \\ \phi_{0,2} + \sum_{h=1}^p \phi_{h,2} Y_{t-h} + \epsilon_t & \text{per } B_1 \leq Z_t < B_2 \\ \dots & \\ \phi_{0,r} + \sum_{h=1}^p \phi_{h,r} Y_{t-h} + \epsilon_t & \text{per } B_{r-1} \leq Z_t \end{cases} ,$$

con  $\epsilon_t \sim WN(0; \sigma^2)$ , indipendente dalla storia passata di  $Y_t$ .

La generalizzazione al caso  $r > 1$  aumenta la dipendenza della serie  $Y_t$  dal processo indipendente, avvicinandolo sostanzialmente ai modelli Autoregressivi a ritardi distribuiti, o *Autoregressive Distributed Lag* (ADL), che inseriscono il regressore esterno nella relazione lineare. Gli ADL non ne costituiscono tuttavia una generalizzazione, perché nei TAR l'effetto della variabile esterna è su tutti i parametri. Si può tuttavia, prendendo spunto dagli ADL o meglio dai modelli di regressione multipla per le serie storiche, definire una ulteriore modifica del modello TAR, che prevede di aggiungere negli  $r$  regimi dei regressori, con il vantaggio, se questi ultimi sono significativi, di aumentare l'adattamento del modello alla serie. Ad esempio, preso per semplicità un solo processo  $X_t$  come regressore e due soli regimi, si può definire il modello Autoregressivo a barriera e a ritardi distribuiti (TADL), con  $p_1$  il ritardo massimo del processo  $Y_t$ ,  $p_2$  il ritardo massimo del processo  $X_t$  e 2 il numero di regimi:

$$Y_t = \begin{cases} \alpha_0 + \sum_{h=1}^{p_1} \alpha_{h,1} Y_{t-h} + \sum_{h=1}^{p_2} \alpha_{h,2} X_{t-h} + \epsilon_t & \text{per } Z_t < B \\ \beta_0 + \sum_{h=1}^{p_1} \beta_{h,1} Y_{t-h} + \sum_{h=1}^{p_2} \beta_{h,2} X_{t-h} + \epsilon_t & \text{per } Z_t \geq B \end{cases} .$$

Come osservato in precedenza, il TAR è un caso generale dei modelli per i *break* strutturali, infatti si possono ottenere i secondi dai primi ponendo  $Z_t = t$  e  $B = \tau$ . La scelta di quale variabile utilizzare per  $Z_t$  è legata alle conoscenze ex ante che si hanno sul processo interno, e può essere non soltanto un processo esogeno, ma bensì anche un processo endogeno, con cui si intende che la stessa variabile dipendente può essere utilizzata come  $Z_t$ , essa richiede però di essere presa con un certo ritardo. Tale modello è il secondo tra quelli che si è scelto di riportare nel presente capitolo, ed è chiamato il *Self Exciting Threshold Autoregressive Model* (SETAR) o Modello a Barriera Autoeccitante :

**Definizione (Modello a Barriera Autoeccitante a due regimi)** *Il Modello a Barriera Autoeccitante a due regimi di ordine uno è definito come un modello TAR(1, 2) per il quale il processo endogeno è pari a:  $Z_t = Y_{t-d}$  con  $d = 1, 2, \dots$ , ovvero:*

$$Y_t = (\alpha_0 + \alpha_1 Y_{t-1})(1 - |Y_{t-d} \geq B|) + (\beta_0 + \beta_1 Y_{t-1})|Y_{t-d} \geq B| + \epsilon_t .$$

La particolarità dei modelli SETAR si adatta bene al problema dell'analisi del *mass lapse*, perché permette di introdurre quell'effetto di contagio più volte richiamato, in questi termini: quando i riscatti raggiungono una certa soglia, la serie passa ad un regime a media maggiore. Il difetto di questi modelli è però di non poter stimare il differimento  $d$ , che deve essere invece scelto a priori. Una tecnica possibile di selezione prevede di confrontare le diverse stime prodotte dai modelli con diverso differimento  $d$ ; si possono ad esempio confrontare le significatività dei parametri, con il problema però di falsare il risultato dei test, oppure, procedura migliore, confrontare gli AIC o i BIC dei modelli e selezionando quello con il valore più basso.

Passando alla stima dei parametri, un elemento interessante in merito che può essere

dimostrato è che per avere la stazionarietà - e quindi avere tutte le buone proprietà dette in precedenza degli stimatori - non è necessario che i regimi siano a loro volta stazionari, si possono ottenere coefficienti superiori ad uno in valore assoluto e mantenere comunque la stazionarietà. Vi sono due tecniche proposte in letteratura per la stima del modello, qui viene presentata solo la prima, intuitivamente più semplice. L'idea è quella di minimizzare rispetto alla soglia  $B$  la somma al quadrato dei residui (RSS), avendo stimato il vettore dei parametri per ogni valore di  $B$  scelto. Più nel dettaglio, la procedura per la stima di un modello TAR(1, 2) si svolge come segue: si definisce un intervallo reale chiuso a cui appartenga la soglia:  $B \in [B_{min}, B_{max}]$ , supponendo di avere un campione temporale di ampiezza  $T$ , viene riscritto il modello TAR in forma vettoriale. Ciò lo si può fare partendo dalla formula presentata nei modelli per i *break*, avendo svolto le parentesi:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 Y_{t-1} + \beta_2 |Z_t \geq B| + \beta_3 Y_{t-1} |Z_t \geq B| + \epsilon_t = \\ &= \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t}(B) + \beta_3 X_{3,t}(B) + \epsilon_t , \\ \text{dove } X_{1,t} &= Y_{t-1} , \quad X_{2,t}(B) = |Z_t \geq B| \quad \text{e} \quad X_{3,t}(B) = Y_{t-1} |Z_t \geq B| . \end{aligned}$$

Nella formulazione si può riconoscere un semplice modello di Regressione Lineare, dove le ultime due covariate sono in funzione della soglia. Si hanno perciò una famiglia di modelli indicizzati rispetto ai valori di  $B$  e stimabili ognuno secondo i minimi quadrati. In termini matriciali si ottiene la consueta formulazione dei modelli di Regressione Lineare Multipla:

$$\underline{Y} = X(B)\underline{\beta} + \underline{\epsilon}$$

$$\underline{Y} = \begin{bmatrix} Y_2 \\ \dots \\ Y_T \end{bmatrix} , \quad X(B) = \begin{bmatrix} 1 & X_{1,2} & X_{2,2}(B) & X_{3,2}(B) \\ \dots & & & \\ 1 & X_{1,T} & X_{2,T}(B) & X_{3,T}(B) \end{bmatrix} , \quad \underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} , \quad \underline{\epsilon} = \begin{bmatrix} \epsilon_2 \\ \dots \\ \epsilon_T \end{bmatrix} .$$

Fissato un valore di  $B$ , e dati  $\hat{y}$  e  $\hat{X}(B)$  i valori osservati del campione, si può definire la somma al quadrato dei residui:

$$RSS(\underline{\beta}|B) = [\hat{y} - \hat{X}(B)\underline{\beta}]^t [\hat{y} - \hat{X}(B)\underline{\beta}] .$$

Minimizzando la funzione rispetto al vettore dei parametri, per  $B$  fissato, si ottiene la soluzione classica, in funzione però della soglia:

$$\hat{\underline{\beta}} = \hat{\underline{\beta}}(B) = [\hat{X}(B)^t \hat{X}(B)]^{-1} \hat{X}(B)^t \hat{y} .$$

A questo punto, avendo il vettore stimato in funzione della soglia, si minimizzano i RSS rispetto a  $B$ , utilizzando nella formula le stime precedentemente ottenute:

$$RSS(B|\hat{\underline{\beta}}) = [\hat{y} - \hat{X}(B)\hat{\underline{\beta}}(B)]^t [\hat{y} - \hat{X}(B)\hat{\underline{\beta}}(B)] .$$

Il valore  $B$  che minimizza la formula è la stima. Si può dimostrare che gli stimatori dei parametri così costruiti soddisfano le stesse proprietà degli OLS.

## 4.2 I Modelli STAR

Per la modellizzazione di fenomeni come quello studiato dalla presente ricerca, i modelli TAR possono dimostrarsi molto utili, ma carenti su un punto: la transizione tra gli stati. Ciò che si osserva solitamente infatti è una transizione più lenta per fenomeni come quello del *mass lapse*, per il quale il picco è preceduto da una fase di crescita che collega i due regimi. Nei modelli di regressione con *dummy* la transizione è istantanea, mentre, come detto in precedenza, nei modelli TAR la transizione è mediata dal fattore autoregressivo, l'effetto però può comunque risultare non sufficiente. Una scelta potrebbe essere quella di aumentare il numero di regimi, ad esempio stimare un regime intermedio che si adatti alla fase di crescita. Anche così tuttavia si hanno dei problemi, ad esempio la fase di transizione potrebbe dipendere da troppi pochi valori per poter essere stimata correttamente. Soprattutto però si osserva che aumentando il numero dei regimi la complessità computazionale aumenta, rendendo più lunga e meno affidabile la procedura di stima. Per questo motivo si è pensato ad una evoluzione dei TAR, che sorge abbastanza naturalmente dalle formule presentate in precedenza, che permette di avere infiniti regimi, ma facilmente stimabili. Questi sono i cosiddetti modelli *Smooth Transition Autoregressive Model* o modelli a Transizione Graduale (STAR). Per vedere come nasce l'idea di questi modelli si può prendere la formula dei TAR(1,2) espressa tramite la variabile indicatrice:

$$Y_t = (\alpha_0 + \alpha_1 Y_{t-1})(1 - |Z_t \geq B|) + (\beta_0 + \beta_1 Y_{t-1})|Z_t \geq B| + \epsilon_t ;$$

quando il processo  $Z_t$  supera la barriera allora l'indicatore dell'evento risulta pari a uno e nella formula rimane solo il secondo addendo, specularmente, quando  $Z_t$  è sotto la barriera, l'indicatore è pari a zero e rimane il primo addendo. Nella formulazione del modello si può riconoscere una media ponderata di due regimi, con pesi pari a zero e uno, di conseguenza una generalizzazione naturale si può ottenere utilizzando pesi diversi, in particolare sostituendo all'indicatore dell'evento un processo stocastico che dipenda dal processo esterno e dalla soglia e che sia compreso tra zero e uno. Il processo risultante è tale per cui al variare di  $Z_t$  si hanno diversi regimi con parametri le medie ponderate dei parametri dei due regimi principali della formula. I modelli che fanno uso di questa generalizzazione sono appunto i modelli STAR, i quali possono essere adesso meglio definiti:

**Definizione (Modelli Autoregressivi a Transizione Graduale)** Il processo  $\{Y_t; t \in \mathbb{Z}\}$  segue un Modello Autoregressivo a Transizione Graduale di ordine 1, STAR(1) se e solo se vale:

$$Y_t = (\alpha_0 + \alpha_1 Y_{t-1})(1 - G(Z_t; \gamma, B)) + (\beta_0 + \beta_1 Y_{t-1})G(Z_t; \gamma, B) + \epsilon_t ,$$

dove  $\epsilon_t \sim WN(0, \sigma^2)$ , indipendente dalla storia passati di  $Y_t$ , e dove  $G(Z_t; \gamma, B)$  è una trasformata deterministica del processo endogeno  $\{Z_t; t \in \mathbb{Z}\}$ , tramite la funzione  $G(x; \gamma, B)$ , chiamata “funzione di transizione”, tale che:  $G : \mathbb{R} \rightarrow [0, 1]$ , dipendente dal parametro soglia  $B \in \mathbb{R}$  e dal parametro di gradazione  $\gamma$ .

I due regimi assunti dal processo sono indicati con i termini “regime basso” e “regime alto”, dato che solitamente i due regimi hanno media differente, anche se di poco<sup>93</sup>.

La funzione  $G$  ammette diverse forme, tra le quali anche quella dell’indicatore dell’evento  $|Z_t \geq B|$ , che permette di ottenere il modello TAR come caso particolare. Più interessanti sono le funzioni continue, per le quali vale che la transizione dipende dalla distanza tra il processo  $Z_t$  e la barriera. Un esempio potrebbe essere quello di una funzione per la quale se il processo indipendente supera  $B$  ci si sposta verso il regime alto, e specularmente, se scende sotto  $B$ , verso il regime basso. Per quanto riguarda il parametro di gradazione  $\gamma$ , esso serve invece a controllare la rapidità della transizione.

Passando alle possibile forme funzionali di  $G$ , vengono solitamente proposte due scelte: una funzione esponenziale, il cui modello corrispondente è chiamato ESTAR, e la funzione logistica, il cui modello corrispondente viene invece chiamato LSTAR:

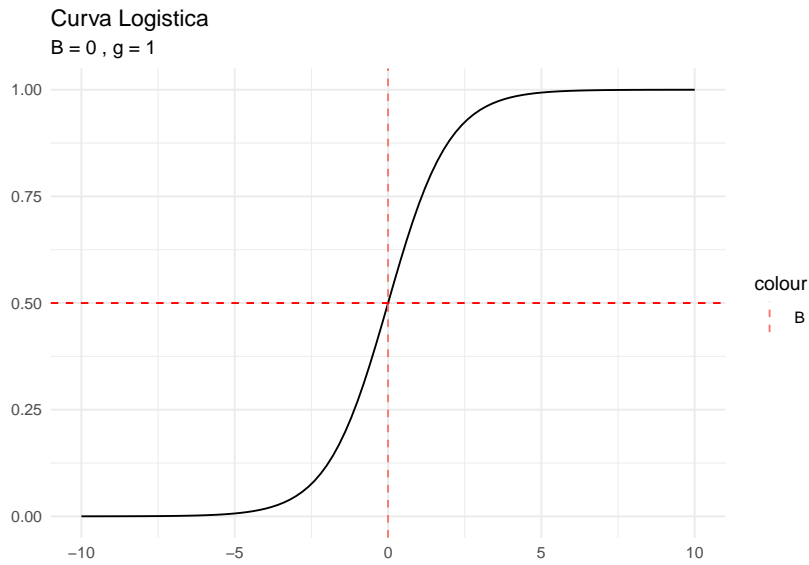
**La funzione logistica del modello LSTAR** Partendo dal modello LSTAR, la funzione utilizzata come detto è la logistica, espressa però nei due soli parametri dati in precedenza, dato che è possibile esprimere la curva in maniera più generale:

$$G(x; \gamma, B) = \frac{1}{1 + e^{-\gamma(x-B)}} , x \in \mathbb{R}.$$

Per una sua comprensione migliore viene riportato il grafico della funzione per i valori  $B = 0$  e  $\gamma = 1$ :

---

<sup>93</sup>Si potrebbe avere il caso di intercetta costante uguale, mentre il parametro associato alla componente autoregressiva diverso; sarebbero sempre un modello STAR, ma è un caso meno interessante per la presente Tesi.



Si può dimostrare tramite uno studio di funzione che la logistica è strettamente monotona, crescente per  $\gamma > 0$ , decrescente per  $\gamma < 0$ , il punto di flesso si sposta per diversi valori di  $B$ , mentre i suoi limiti sono:

- se  $x \rightarrow -\infty$  allora  $G \rightarrow 0$ ;
- se  $x \rightarrow +\infty$  allora  $G \rightarrow 1$ ;
- se  $\gamma \rightarrow +\infty$  allora  $G \rightarrow |Z_t \geq B|$ , che implica che il modello LSTAR diventa un modello TAR;
- se  $\gamma \rightarrow 0$  allora  $G \rightarrow 0,5$ , che implica che il modello LSTAR diventa un modello AR con parametri pesati dai due regimi, pesi pari a 0,5.

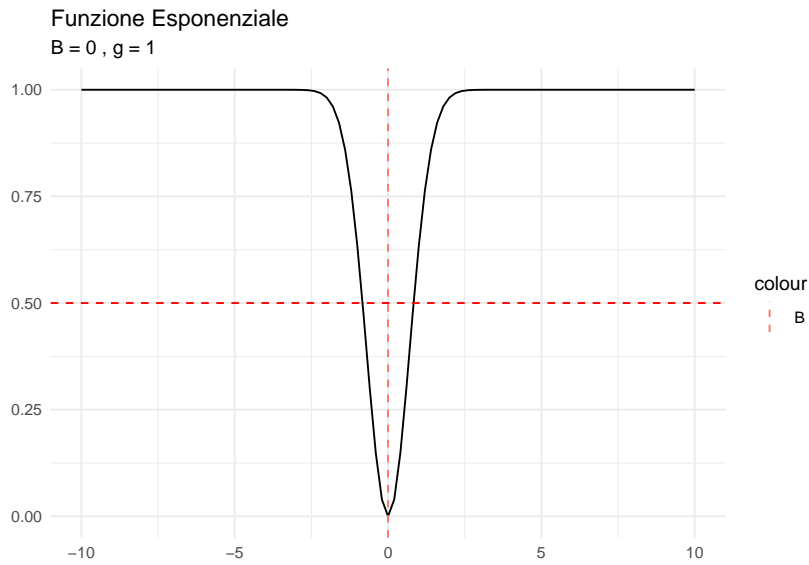
Il modello che risulta dall'utilizzo di questa funzione permette di descrivere il comportamento di una serie per la quale più il processo indipendente si allontana dalla soglia più il processo dipendente si sposta verso uno dei due regimi, senza mai raggiungerli perfettamente.

**La funzione esponenziale del modello ESTAR** Per quanto riguarda invece il modello ESTAR, la funzione che la caratterizza è una particolare funzione esponenziale così costruita:

$$G(x; \gamma, B) = 1 - e^{-\gamma(x-B)^2}, \text{ per } \gamma > 0,$$

nella cui forma funzionale si può riconoscere il nucleo della densità di una *Normale*. Qui di seguito si può osservare il suo grafico per  $\gamma = 1$  e  $B = 0$ . Si può osservare appunto la classica curva a “campana” della *Normale*, in questo caso però rovesciata, a causa del segno meno davanti all'esponenziale:





Dallo studio di funzione si ricavano le seguenti proprietà:

- se  $x \rightarrow \pm\infty$  allora  $G \rightarrow 1$ ;
- se  $\gamma \rightarrow 0$  allora  $G \rightarrow 0$  e il modello converge ad un AR;
- se  $\gamma \rightarrow +\infty$  allora  $G \rightarrow 0$  e il modello converge ancora ad un AR.

In questo caso il modello ESTAR serve a descrivere una serie per la quale più  $Z_t$  si avvicina alla soglia e più il processo tende ad un certo regime, mentre quando  $Z_t$  si allontana da  $B$ , sia da sotto che da sopra, si avvicina al secondo regime.

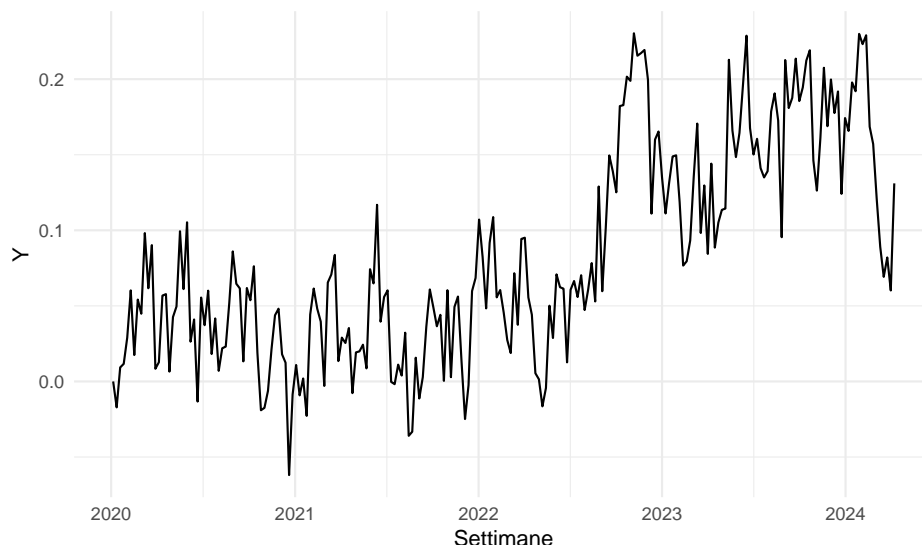
Parlando del processo indipendente  $Z_t$ , anche in questo caso la variabile può essere scelta in diversi modi: un processo esogeno ritenuto significativo nel determinare il salto, la variabile di indagine stessa, presa in un suo ritardo, ottenendo l'equivalente STAR dei modelli SETAR, anche detti SESTAR: *Self Exciting Smooth Transition Autoregressive models*, e in fine si può prendere il tempo come variabile e come soglia una certa data significativa.

Anche per questo modello viene riportato di seguito un esempio simulato per comprendere meglio il comportamento degli STAR. Si è scelto un modello LSTAR(1, 2) con parametri e variabile sottostante come nel caso dell'esempio per i modelli TAR, tranne per una varianza ridotta a 0.001 per poter meglio osservare i due regimi. La funzione  $G$  scelta è la logistica con parametri  $B = 0.035$  e  $\gamma = 80$ .

### Simulazione di un processo LSTAR(1,2)

Parametri:  $a=(0;0.4), b=(0,1;0,7), s_2 = 0,001, B=0,035$  e  $g=80$ .

Variabile indipendente: BTP rendimenti settimanali, 2020/01/01–2024/04/07.



Per quanto semplice sia la costruzione del modello, esso presenta forti criticità, in particolare le condizioni per la stazionarietà sono molto complesse e ancora non pienamente note. La stima dei parametri è operata tramite una tecnica che prende il nome di *Non Linear Least Squares estimation* (NLS), ovvero si minimizzano gli errori al quadrato del modello, così come accade per gli OLS, ma poiché i parametri non sono lineari nel modello, non è possibile ottenere in generale un risultato in forma chiusa, perciò la minimizzazione è svolta tramite un algoritmo numerico. Gli stimatori ottenuti, se valgono le usuali ipotesi, soddisfano i criteri di correttezza, consistenza e asintotica normalità. Tra le ipotesi è richiesto che il processo sia autocorrelato solo fino ad un certo ritardo, per avere l'indipendenza asintotica. La procedura di per sé non è esente da errori, il massimo raggiunto potrebbe essere solamente un ottimo locale, e ciò dipende fortemente dalla scelta delle condizioni per inizializzare l'algoritmo. Inoltre, la procedura applicata ai modelli STAR ha difficoltà di stima per il parametro  $\gamma$  quando questo risulta elevato e la funzione scelta è la logistica, dato che, per valori grandi del parametro, il modello diventa difficilmente distinguibile dai TAR<sup>94</sup>.

I modelli STAR possono essere ulteriormente estesi per ammettere più soglie, mantenendo sempre due regimi, in modo tale da modificare l'andamento della transizione e ottenere forme più particolari per la funzione  $G$ . L'estensione si può applicare alla curva Logistica in questo modo: dati  $B_1 < \dots < B_n$ , la funzione logistica a più soglie è definita come:

$$G(x; \gamma, B_1, \dots, B_n) = \frac{1}{1 + e^{-\gamma(x-B_1)\dots(x-B_n)}} .$$

Una ulteriore modifica permette di introdurre in aggiunta altri regimi. Entrambe queste generalizzazioni presentano nuove problematiche, in particolare per la stima e per

<sup>94</sup>Per una sintesi della stima dei modelli STAR si può fare riferimento a Krolzig (2002).

l'analisi del comportamento della serie, inoltre non si mostrano in generale necessarie poiché il risultato è difficilmente distinguibile dai casi più semplici; per questo motivo non saranno qui affrontate né utilizzate.

### 4.3 Il “*Modelling Cycle*” per i modelli STAR

Questi particolari modelli sono meglio riassunti nell'articolo Dijk, Teräsvirta, and Franses (2002). In esso gli autori riportano un suggerimento di “*Modelling Cycle*”, ovvero, una lista di procedure da seguire per l'applicazione e la diagnostica dei modelli STAR. Ecco come si presentano i diversi punti:

1. Specificare un modello AR di ordine  $p$  per il processo indagato;
2. Verificare tramite test apposito se la serie è lineare, contro l'ipotesi di modello STAR o TAR. Nel caso in cui sia valutata non lineare, selezionare l'apposito processo indipendente;
3. Stimare i parametri del modello STAR;
4. Eseguire test e diagnostica per verificare la bontà del modello;
5. Modificare il modello in modo opportuno rispetto al punto 4;
6. Usare il modello risultante per una analisi descrittiva o per fare previsioni.

Mentre nei paragrafi precedenti sono stati indagati il primo e il terzo punto, si intende ora soffermarsi sul secondo e sul quarto, riguardanti la verifica di ipotesi per la validazione del modello. Andando nel dettaglio, il secondo punto ha come scopo di verificare se la serie è lineare, e quindi rappresentabile tramite un semplice modello AR, o se invece è non lineare, e in quel caso verificare se i modelli STAR o TAR possano essere più adeguati per la serie rispetto all'AR. Il quarto invece si concentra sulla verifica delle ipotesi del modello scelto e sulla valutazione dell'adattamento del modello scelto, tra STAR e TAR, ai dati osservati.

Il problema che affligge l'inferenza statistica per questa classe di modelli è il fatto che le statistiche convenzionali non possono essere usate, poiché i test hanno distribuzioni asintotiche non standard, ovvero la distribuzione a cui converge è non nota e non esprimibile analiticamente. Un caso specifico esemplificativo è il test di linearità del modello contro l'ipotesi di un modello STAR, dove si verifica che sotto ipotesi nulla il modello presenta parametri non identificabili. Infatti, dato un modello STAR(1), le due ipotesi - nulla e alternativa - possono essere espresse sotto due formulazioni logicamente equivalenti:

$$\left\{ \begin{array}{l} H_0 : (\alpha_0, \alpha_1) = (\beta_0, \beta_1) \\ H_1 : \alpha_0 \neq \beta_0 \vee \alpha_1 \neq \beta_1 \end{array} \right. \quad \text{o} \quad \left\{ \begin{array}{l} H_0 : \gamma = 0 \\ H_1 : \gamma \neq 0 \end{array} \right. .$$

Nello stimare il modello sotto ipotesi nulla, si ha che la verosomiglianza si semplifica rispetto ai parametri non considerati: nella prima formulazione si elidono i parametri  $\gamma$  e  $B$ , nella seconda i parametri  $\alpha$  e  $\beta$ . Il test risultante presenta una distribuzione asintotica non standard non determinabile, perciò si ricorre tipicamente a due soluzioni: la prima prevede di ottenere in via simulativa alcuni valori critici della distribuzione, mentre la seconda sfrutta l'approssimazione tramite polinomio di Taylor della funzione  $G$ , che permette di risolvere il problema di identificazione e di ottenere dei risultati asintotici noti; in particolare, le statistiche ottenibili sono le classiche a distribuzione  $\chi^2$  o  $F$ <sup>95</sup>.

Detto ciò, i test che vengono costruiti per il secondo punto sono tipicamente il test di linearità contro un modello SETAR, contro un modello LSTAR e contro un modello ESTAR; questi ultimi due utilizzano l'approssimazione tramite polinomio di Taylor. Passando invece al quarto punto, alcuni test rimangono gli stessi dei modelli lineari ordinari, come i test per la *Normalità* dei residui - un'ipotesi in generale non necessaria per la stima di questi ma che può risultare utile per fare previsioni - , o ancora i test di significatività, che rimangono equivalenti e valgono però solo asintoticamente, e i test per l'autocorrelazione dei residui. Un test specifico invece per la verifica dei modelli TAR o STAR è il test di ipotesi di "assenza di non linearità residua": l'intento è quello di verificare se il modello risultante presenti ancora una non linearità residua, in particolare valutando se il numero di regimi è sufficiente a descrivere il comportamento della serie. Il test si configura pertanto in modo simile ai test di confronto del secondo punto, dove però in questo caso il confronto è svolto tra un modello TAR a regime  $p$  e modello TAR a regime  $q > p$ , e congruamente per i STAR. Il test soffre dello stesso problema di parametri non identificabili e si ricorre al metodo già citato di approssimazione tramite polinomio di Taylor.

#### 4.4 Un modello con *trend* deterministico non lineare

I modelli TAR, SETAR e STAR costituiscono la parte principale dei modelli utilizzati dalla presente ricerca per l'analisi del tasso di riscatto. Si è scelto però di aggiungere a questi un ulteriore approccio, tecnicamente meno avanzato ma più immediato nella sua applicazione e interpretazione, basato sull'idea dei modelli per le serie storiche con *trend* non lineare deterministico. Il modello è costituito da una componente autoregressiva, più una funzione del tempo. Tale funzione rappresenta questo *trend* o andamento di fondo, il quale, con riferimento al caso sotto esame, serve a riprodurre il comportamento previsto, secondo quanto espresso dalla letteratura, della serie dei riscatti in presenza dell'evento di massa.

---

<sup>95</sup>Sui test statistici per gli STAR si può fare riferimento a Dijk, Teräsvirta, and Franses (2002).

Andando più nello specifico, le ipotesi che accompagnano il modello e che descrivono la serie dei riscatti sono così espresse: i riscatti mantengono un andamento lineare, autoregressivo e stazionario, nei periodi in cui non è presente l'evento di massa; nell'istante in cui si verifica l'evento, accade la rottura della stazionarietà: da un *trend* costante si passa ad un altro *trend*, descritto da una crescita rapida dei tassi fino ad un massimo, e poi una decrescita lieve fino al ritorno al regime stazionario. Pertanto la funzione ideale  $f(t)$  che rappresenta il *trend* deterministico, con  $t \in \mathbb{Z}$ , deve possedere queste caratteristiche:

1.  $f(t) \geq 0$ , per ogni  $t \in \mathbb{Z}$ ;
2. dato l'istante  $S \in \mathbb{Z}$  in cui si verifica l'evento di massa, allora  $f(t) = 0$  per  $t < S$ ;
3. la funzione ammette un massimo in  $t^* \geq S$ ,  $t^* \in \mathbb{Z}$ ;
4.  $\lim_{t \rightarrow \infty} f(t) = 0$ .

Le ipotesi fornite implicano che si stia analizzando la serie dei riscatti in un intervallo di tempo ristretto nel quale è possibile si verifichi un solo *mass lapse*. Il quarto punto impone che l'effetto del *trend*, causato dal *mass lapse*, si esaurisce nel tempo.

Il modello in generale si può esprimere come segue:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + f(t) + \epsilon_t$$

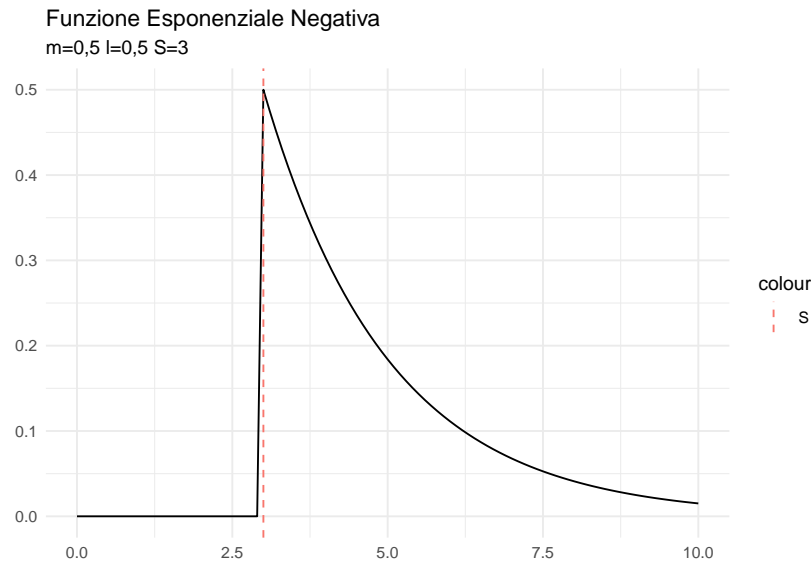
Come in precedenza, si suppone che  $\epsilon_t$  sia una *White Noise* indipendente dalla storia passata di  $Y_t$ . Si suppone, inoltre, che la funzione  $f$  dipenda da dei parametri che si intende stimare dal campione; tali parametri potranno essere non lineari rispetto a  $t$ . La stima del modello dovrà perciò essere operata tramite il metodo già in precedenza citata dei *Non Linear Least Squares estimation*. La procedura, si può ricordare, permette di ottenere delle stime corrette, consistenti e asintoticamente *Normali*, tuttavia presuppone che l'istante  $S$  sia dato, o meglio, è possibile ottenere una stima dell'istante, ma in generale non sono assicurate le buone proprietà degli stimatori sopraelencate. La procedura richiede quindi di determinare in altra maniera, esogenamente, l'istante di *mass lapse*.

Di seguito vengono proposte due curve che possono essere utilizzate per descrivere l'andamento del rischio in esame:

**Curva esponenziale negativa** Il primo esempio si rifà alla funzione esponenziale negativa, con due parametri, definita in questo modo:

$$f(t) = \begin{cases} 0 & \text{per } t < S \\ \mu e^{-\lambda(t-S)} & \text{per } t \geq S \end{cases}, \quad \mu, \lambda > 0.$$

Il grafico è il seguente:



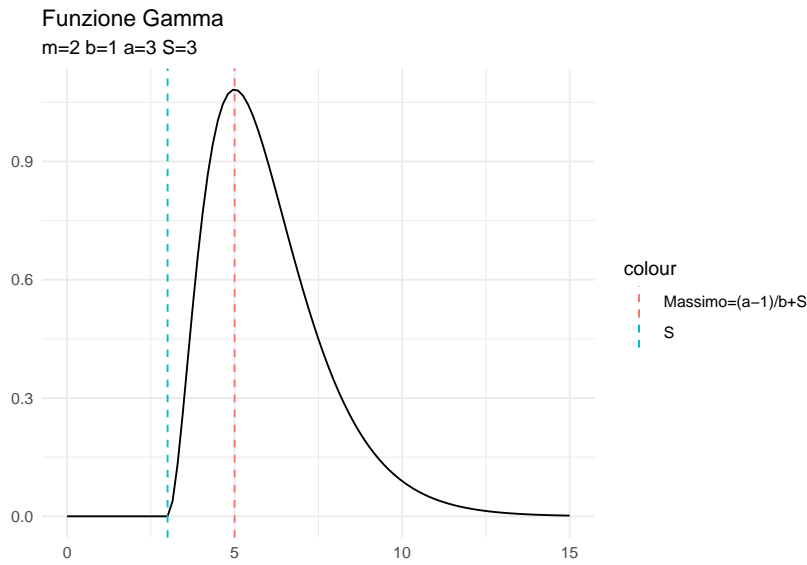
La curva riprende l'idea della funzione di risposta ad un impulso di un AR, l'ampiezza dell'effetto però non è in questo caso data dal valore stocastico dell'impulso  $\epsilon_t$  ma dal parametro certo  $\mu$ , che quindi governa l'ampiezza del salto dei riscatti.

La funzione ha il difetto di avere un impatto immediato. Se si fosse interessati a rappresentare il fenomeno con una crescita iniziale più lenta, si potrebbe passare ad una curva più a campana, come una *Gaussiana*. Pure quest'ultima presenta un difetto: è simmetrica, ovvero sia, la parte crescente e la parte decrescente, corrispondenti rispettivamente alla fase ascendente e alla fase calante dei riscatti, sono speculari. Una forma più realistica ipotizzabile è quella che rappresenta il fenomeno con una curva che cresca più rapidamente rispetto a quanto cala. Una curva simile la si può ottenere ricorrendo alla densità di una *Gamma* traslata:

**Curva gamma** La funzione, per i valori di  $t$  superiori a  $S$ , riprende il nucleo della densità della *Gamma* traslata:

$$f(t) = \begin{cases} 0 & \text{per } t < S \\ \mu e^{-\beta(t-S)}(t-S)^{\alpha-1} & \text{per } t \geq S \end{cases}, \mu, \beta, \alpha > 0.$$

Per  $\alpha \geq 1$  la curva presenta un massimo, in particolare, per  $\alpha = 1$  si ottiene la curva esponenziale, che ha massimo in  $S$ . Nel caso di  $0 < \alpha < 1$ , la funzione presenta un asintoto verticale in  $S$ , pertanto la curva gamma non è sempre dotata di massimo. Di seguito un grafico della funzione:



Nel grafico è stato indicato con la linea tratteggiata rossa l'ascissa del punto di massimo della funzione. Per la determinazione del punto di massimo si rimanda alle Appendici.

Tale curva è stata scelta per l'analisi numerica successiva. Il modello risultante può essere riscritto in forma compatta nel modo seguente<sup>96</sup>:

$$\begin{aligned}
 Y_t &= \beta_0 + \beta_1 Y_{t-1} + \beta_2 e^{-\gamma_1 Z_t} Z_t^{\gamma_2} + \epsilon_t \\
 Z_t &= (t - S)_+
 \end{aligned}$$

Il modello può essere modificato introducendo altri regressori come quelli già in precedenza elencati. Inoltre, si può osservare che il modello è di fatto equivalente ad un modello TAR per il quale la variabile esterna è il tempo e il secondo regime presenta un *trend* non lineare nel tempo. Allora, delle ragionevoli modifiche potrebbero essere quelle di sostituire la variabile esterna utilizzata o di stimare il parametro  $S$  con il metodo visto in precedenza per i modelli TAR.

Si è scelto di introdurre questo ulteriore modello perché si adatta molto bene nel rappresentare l'andamento della serie nel momento in cui si osserva il picco estremo. Di per sé non ha valenza predittiva dato che: non descrive il comportamento dei riscatti nella sua totalità, non studia le cause del fenomeno, ma soprattutto suppone che si verifichi una solo picco, quando invece ci si attende che la serie dei riscatti subirà in futuro ulteriori salti. Permette però di stimare, nel breve, un incremento ragionevole e la decrescita successiva e si può usare perciò come modello semplice di confronto per gli RSM. L'approccio si ispira ad alcuni tipi di modelli statistici che sono stati utilizzati anche durante il periodo del *Covid-19* per analizzare l'andamento del picco

<sup>96</sup>I nomi delle variabili sono stati riscritti per essere coerenti con i simboli tipici dei modelli di regressione

della mortalità osservata. Un esempio che si può riportare è in Valvo (2020) che utilizza una curva bimodale per descriverne il *trend* dei decessi in Italia.

## 4.5 Le critiche ai modelli

A conclusione del paragrafo, dopo aver presentato i diversi modelli, si vuole qui ora mettere a risalto alcune criticità che si possono muovere a questi. Innanzitutto, partendo dai RSM, per quanto semplici nella loro definizione, essi presentano un notevole numero di complessità, sia per la stima che per la valutazione della stazionarietà e della bontà del modello, e possono quindi essere molto ostici da applicare. Inoltre, la loro complessità aumenta se si aumenta il numero dei regimi e soprattutto se si introducono covariate. Per quanto poi studiati e applicati, attualmente sono poco implementati nei *software* comuni; come ricordato in precedenza su R è presente un solo pacchetto per l'analisi dei modelli STAR.

Oltre a tali problemi, gli RSM presentano anche dei punti problematici per la rappresentazione del tasso di riscatto: l'ipotesi alla base di questi è che il cambio di regime dipenda da una sola variabile esterna, mentre è ragionevole supporre che intercorrano diversi fenomeni a modificarne il corso; non solo, il passaggio dal primo al secondo regime potrebbe dipendere da una variabile diversa rispetto a quella che governa il passaggio in direzione opposta, dal secondo al primo regime. Si può pensare realisticamente che il salto dai riscatti bassi ai riscatti estremi sia dovuto all'aumento dei tassi di mercato, mentre la decrescita sia governata dalle azioni intraprese dalla compagnia atte a ridurre il fenomeno.

Un ulteriore problema è poi il seguente: si può pensare che non sempre l'innalzamento dei tassi provochi l'evento di massa, mentre gli RSM assumono la transizione come certa, in caso il processo indipendente superi la barriera, ed anche la barriera stessa è difficile supporla come un parametro costante. Il fatto che il superamento della barriera possa non produrre il salto tra regimi è legato al problema della stima dell'effetto di correlazione e all'effetto causale. Nell'analisi delle serie storiche si è interessati solitamente a valutare l'effetto causale dei regressori sulla variabile dipendente, per questo i modelli costruiti aggiungono particolari ipotesi, come l'indipendenza della matrice *design* dall'errore. Ha causa della complessità teorica degli RSM, si è scelto di concentrarsi solo sulla valutazione della correlazione tra processo esterno e processo dipendente, tralasciando la stima del secondo effetto, sebbene questo sia rilevante nell'ambito della valutazione dei rischi, com'è questo il caso.

Per quanto riguarda invece il modello a *trend* non lineare, già alcune critiche sono state date, ma se ne può aggiungere un'altra molto importante: la funzione scelta è sempre deterministica, quando piuttosto le serie storiche sono meglio descritte da *trend* stocastici, se presenti. Ciò fa sì che non sia possibile, secondo il modello dato, avere



picchi diversi, se non per la variabilità data dagli errori del processo. Una possibile proposta risolutiva può derivare dall'utilizzo dei modelli di regressione *Bayesiana*, con i quali è possibile imporre la stocasticità dei parametri. La procedura richiede però di definire delle distribuzioni a priori, oltre a richiede l'osservazione di più eventi di massa.

## 5 Analisi empirica

Il capitolo precedente ha introdotto la famiglia dei modelli *Regime Switching* come possibile approccio alla modellizzazione del *lapse rate* e soprattutto come metodo per la valutazione del rischio di estinzione di massa. Sono stati forniti, in particolare, i motivi teorici che rendono ragionevole suddetta proposta. Essa richiede, tuttavia, anche una validazione empirica della sua capacità effettiva di rappresentare il fenomeno, per poter giustificare in maniera più completa il suo utilizzo. In questo ultimo capitolo si intende quindi svolgere tale verifica su un campione di dati osservati. Dopo la valutazione statistica della bontà dei modelli scelti, questi verranno utilizzati per poter ottenere delle stime dell’impatto del *mass lapse risk* tramite il VaR, confrontando poi i risultati con il numero previsto dalla formula standard.

### 5.1 Il campione

Il campione è stato fornito dallo Studio Attuariale “*Visintin e Associati*” (SAVA S.r.l.) ed è composto dagli indici di riscatto settimanali, annualizzati, osservati per tre HRG di una compagnia assicurativa italiana<sup>97 98</sup>. Nello specifico, i tre portafogli si riferiscono al medesimo gruppo di polizze di tipo rivalutabile, con tasso accreditato dello 0%. La finestra di tempo di osservazione delle tre serie è composta dalle prime 39 settimane - sulle circa 52 - del 2023. Si è scelto di escludere la prima osservazione di ognuna delle tre HRG dal *dataset* utilizzato per il modello - riducendo il campione a 38 valori per serie - poiché il suo valore è condizionato dal fatto che nei primi giorni dell’anno, a causa delle festività, non è possibile il riscatto. L’impossibilità al riscatto rende il valore osservato di molto inferiore rispetto a quelli relativi alle settimane successive. Ciò lo si può verificare direttamente dal grafico della serie differenziata dell’indice di riscatto per la prima HRG riportato in figura 9; questo comportamento si osserva essere equivalente anche per le altre due HRG.

Si può pensare alle tre serie come soggette ad un fenomeno di stagionalità, che richiede di essere rimosso per poter analizzare correttamente la serie<sup>99</sup>.

---

<sup>97</sup>Si veda il terzo capitolo per la definizione del tasso.

<sup>98</sup>Si ricorda che con HRG si intende un “gruppo di rischi omogeneo”, ovvero un insieme di contratti raggruppati assieme poiché la loro rischiosità, e nello specifico la loro distribuzione, è stata valutata come molto simile. Si veda sempre il terzo capitolo.

<sup>99</sup>Si è scelta la rimozione dell’osservazione e non la modifica poiché non sono state trovate informazioni aggiuntive, come le settimane precedenti la prima del 2023, che potessero fornire delle stime sensate per ricavare un valore destagionalizzato.

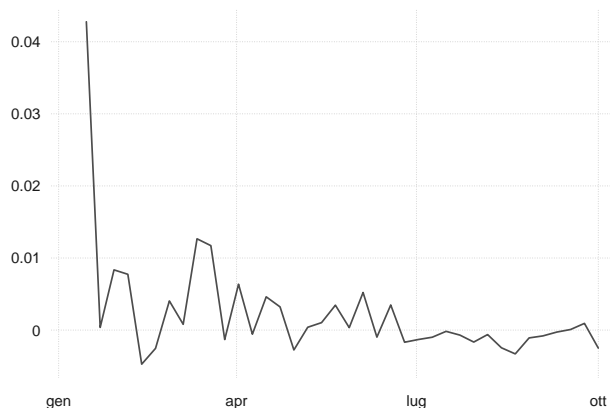


Figura 9: Serie differenziata dell'indice dei riscatti della HRG I.

In figura 10 sono riportati assieme i grafici delle tre HRG, private della prima settimana.

I tre portafogli sono caratterizzati da tre *range* molto distinti: la prima serie va da un circa 6% a un 10%, la seconda invece da un 4% a un 7% e la terza da un 15% a un 35%. In base ai tre range distinti, è possibile ipotizzare per le tre HRG diversi livelli di dinamicità, dalla seconda meno dinamica, alla terza più dinamica<sup>100</sup>. Non si è potuto reperire valori per date precedenti il 2023. Ciò che si è potuto conoscere, da comunicazione con lo studio attuariale, è però che le serie erano più o meno stazionare, poco sotto la media dei primi giorni del 2023; ad esempio la prima HRG doveva aggirarsi in media attorno al 5%. Avendo in mente ciò, si può dire che le tre serie presentano le caratteristiche di uno shock improvviso per i tassi di riscatto, attribuibile appunto ad una rottura o salto tra regimi causato da un evento *mass lapse*.

Un elemento fondamentale da sottolineare è che il *lapse rate* per i tre portafogli è calcolato tramite l'indice di riscatto, definito precedentemente nel terzo capitolo. Si può ricordare che l'indice è basato su quantità monetarie - in particolare l'importo totale pagato per i riscatti in una data epoca - e tale caratteristica lo rende più naturalmente modellizzabile tramite le serie storiche economiche; pertanto l'utilizzo degli RSM si adatta bene all'analisi. Le tre serie sono però troppo corte per poter ottenere delle stime sensate dei modelli e il picco, sebbene presente, è interrotto. Sul suo comportamento successivo è noto soltanto che verso l'inizio del 2024 si è abbassato tornando verso i livelli precedenti il 2023. Per poter produrre un'analisi migliore si è dovuto ricorrere a delle fonti esterne per poter "allungare" la serie ed ottenere un campione significativo. Questo tipo di operazione ha come conseguenza di produrre un campione solo parzialmente reale, pertanto in parte soggettivo, poiché dipendente dalle scelte fatte per la sua costruzione. Le critiche al nuovo campione saranno presentate dopo la sua costruzione.

<sup>100</sup>Per il termine "dinamico", si rimanda al secondo capitolo, ultimo paragrafo.

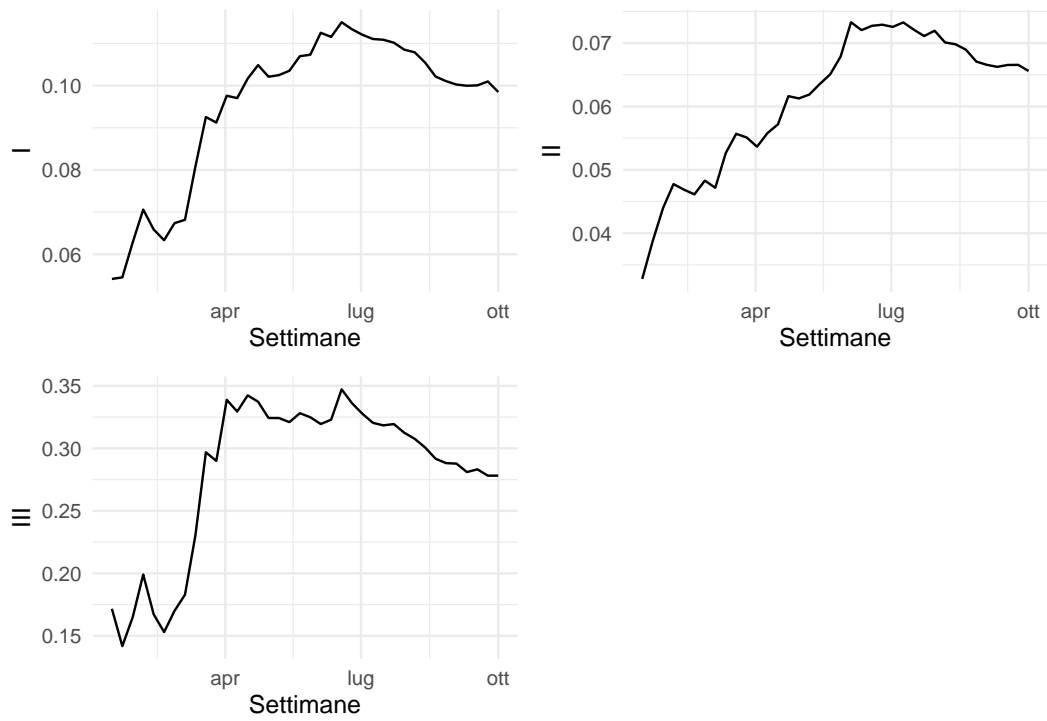


Figura 10: I grafici dell'indice di riscatto osservato per le tre HRG del campione, prive del valore per la prima settimana.

Si può intanto giustificare tale approccio ricordando che il Legislatore consente alle imprese di utilizzare, nelle proprie valutazioni dei requisiti, dati provenienti dal mercato, in assenza di dati interni<sup>101</sup>.

Poiché le imprese non sono in generale intenzionate a rendere pubblici dati personali, soprattutto riguardanti argomenti sensibili, come il numero di riscatti, si è ricorso a dati medi di mercati. Questi sono stati reperiti dal sito dell'ANIA, che presenta un *database*, in parte consultabile liberamente, comprendente diversi valori medi di mercato di compagnie assicurative italiane<sup>102</sup>. I dati riguardanti la serie dei riscatti sono stati ottenuti a partire dal documento, già in precedenza citato, di ANIA: “*ANIA TREND Flussi e Riserve Al IV Trimestre 2023*” (Stefano Trionfetti 2024). Il *paper* di ANIA fornisce, nell'ultima pagina, l'indice di riscatto trimestrale, annualizzato, dal 2020, primo trimestre, al 2023, quarto trimestre. Viene anche riportato, in una pagina precedente, un grafico, sempre dell'indice, a livello però mensile e non annualizzato, dal quale si possono ricavare alcune importanti informazioni, utili per inferire sull'andamento del campione presentato nelle date antecedenti.

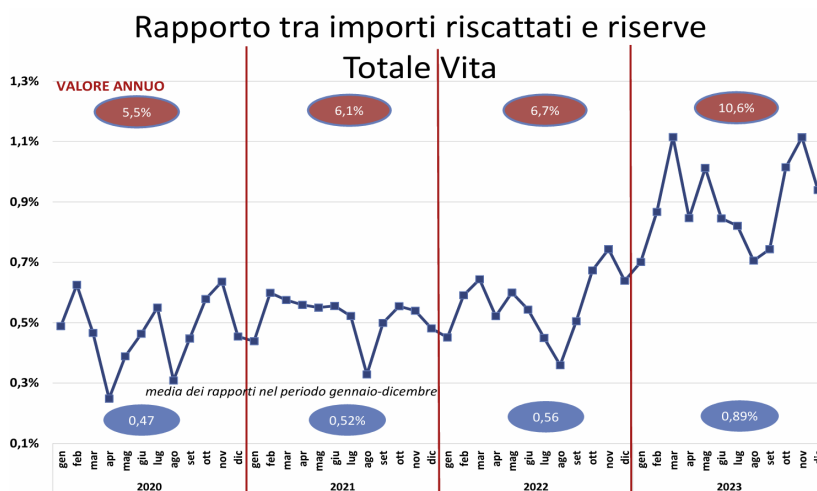


Figura 11: L'indice mensile di riscatto, dal 2020 al 2023. Fonte: Stefano Trionfetti (2024).

Dal grafico riportato in figura 11 si può constatare che il tasso mensile è rimasto più o meno stabile e moderato fino a circa metà 2022, dopo il quale è incrementato fino a raggiungere un secondo regime, caratterizzato da una media più alta dei tassi; quindi l'intero settore italiano è stato colpito da uno *shock* improvviso dei riscatti.

Il particolare *trend* osservato permette, come detto, di inferire sull'andamento del campione in esame nel periodo precedente il 2023. Nello specifico si può confermare,

<sup>101</sup>Si veda il primo capitolo, al paragrafo riguardante il requisito di capitale.

<sup>102</sup>Si riporta il sito dell'ANIA, alla pagina dove poter trovare i *dataset* disponibili al pubblico: [https://www.ania.it/pubblicazioni/-/categories/52472?p\\_r\\_p\\_resetCur=true&p\\_r\\_p\\_categoryId=52472](https://www.ania.it/pubblicazioni/-/categories/52472?p_r_p_resetCur=true&p_r_p_categoryId=52472).

supponendo che il campione sia allineato col mercato, ciò che era stato detto precedentemente, ovvero che le tre serie si aggiravano attorno a valori medi più bassi rispetto al 2023 e che in quell'anno si è verificato uno *shock*.

Come accennato, i dati del grafico non sono disponibili, viene fornito solamente l'indice dei riscatti con frequenza trimestrale, annualizzato, di cui si riporta il grafico in figura 12. Commenti simili a quelli precedenti potrebbero essere fatti per tale campione di dati.

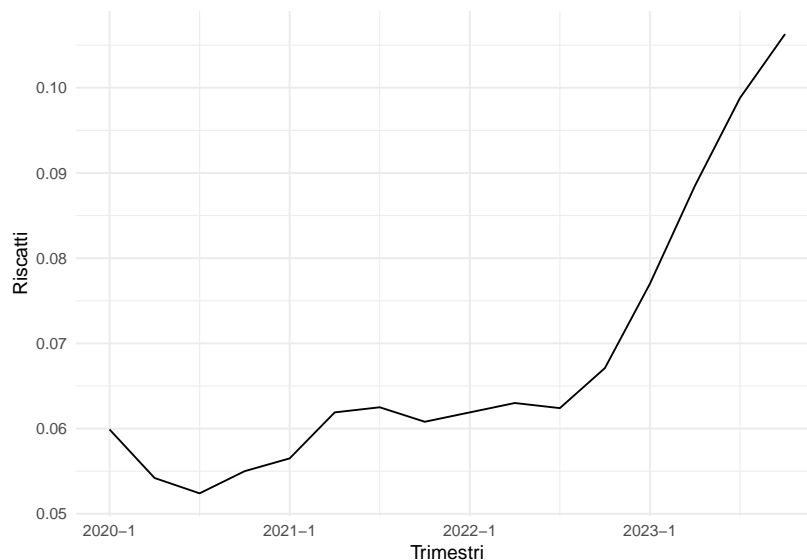


Figura 12: L'indice di riscatto trimestrale per il settore assicurativo italiano. Fonte: Stefano Trionfetti (2024).

Tramite questa serie si vuole poter estendere il campione originale. Sorgono tuttavia due problemi: il primo è che essa ha una frequenza inferiore rispetto alle tre serie - da trimestrale a settimanale - e il secondo è che le dimensioni dei due campioni sono differenti. Con “dimensioni differenti” non si intende semplicemente che le serie hanno *range* diversi, ma che hanno anche una evoluzione temporale diversa: supponendo che esista un istante  $T$  di tempo a cui è legato l'evento di massa, tale istante è ragionevole pensarlo traslato per le due serie. Un fatto importante da osservare è che la serie riportata da ANIA è anch'essa calcolata, come per le tre HRG, tramite l'indice di riscatto. Tale elemento è un requisito necessario per la costruzione del nuovo campione, poiché la scala di misura è la medesima ed è possibile quindi confrontare i valori.

Riassumendo, i due punti cruciali che richiedono ipotesi esterne e che saranno affrontati di seguito sono:

1. La disgregazione temporale dell'indice trimestrale;
2. Il ridimensionamento e la congiunzione dei campioni.

**1) La Disgregazione Temporale:** Esistono diversi approcci alla disgregazione temporale di una serie, in generale però richiedono, per avere buoni risultati, di utilizzare

una serie esterna per la quale si possa supporre il medesimo comportamento, nelle date non osservate, della serie originale<sup>103</sup>. In questo caso non si dispone di una variabile simile, che possa svolgere tale ruolo, e si è ricorso perciò ad un metodo simulativo, più semplice ma meno valido sul piano teorico: simulare per ogni trimestre una *Normale* di media l'indice di riscatto trimestrale e come varianza una specifica varianza stimata per le tre serie. Si sta supponendo implicitamente che i riscatti si siano verificati uniformemente nei trimestri e che la variabilità sia rimasta costante, oltre ad assumere una distribuzione *Normale* per gli errori. Dal punto di vista matematico, l'approccio può essere formalizzato nel modo seguente: approssimando l'anno a 52 settimane, si hanno 13 settimane per trimestre ( $52/4 = 13$ ), dato l'indice di riscatto trimestrale osservato  $\{x_\tau; \tau = 20T1, \dots, 23T4\}$ , la serie simulata per dati settimanali  $x_t^*$  si ottiene in questo modo<sup>104</sup>:

$$x_t^* = \mu_t + \epsilon_t, \quad t = 20S1, \dots, 23S52 ;$$

$$\mu_t = \begin{cases} x_{20T1} & \text{per } t = 20S1, \dots, 20S13 \\ \dots & \\ x_{23T4} & \text{per } t = 23S39, \dots, 23S52 \end{cases} ;$$

$$\epsilon_t \text{ iid } \sim N(0, \sigma^2) .$$

Cruciale diventa la scelta della varianza  $\sigma^2$ : essa deve poter rappresentare la variabilità osservata nelle tre serie e deve perciò venir stimata dal campione. Utilizzare la varianza empirica non sarebbe sensato dato che la unità non possono essere supposte indipendenti, soprattutto considerando che il campione presenta la transizione tra regimi. Si deve poter associare un modello ai dati per poter avere una stima sensata e, data la natura di serie storica, si può ipotizzare per semplicità che le HRG seguano un processo Autoregressivo. La scelta di tale ipotesi è guidata da questo ragionamento: è ragionevole supporre che un AR non sia in grado di rappresentare la serie, visto tutto quello che è stato detto in precedenza sui riscatti, e, non solo, è molto probabile che le stime ricavabili siano poco affidabili, data la numerosità del campione; ciò detto, l'intento è ottenere non una stima statistica ma un valore più o meno realistico per simulare la serie, e la scelta dell'AR - che può sembrare in contraddizione rispetto all'intento del capitolo di verificare che la serie segua un RSM - è dettata dal fatto che è una approssimazione di quelli che saranno i modelli usati successivamente. L'ordine dell'AR può essere ricavato dai grafici dell'ACF e della PACF<sup>105</sup>. Per brevità non vengono mostrati i sei grafici, ma tutte e tre le serie sono compatibili con un AR(1).

Si riporta in figura 13, come esempio, la serie simulata per la prima HRG, avendo stimato la varianza secondo il metodo descritto:

---

<sup>103</sup>Sulle tecniche di disgregazione temporale delle serie storiche si faccia riferimento a Sax and Steiner (2013) e alla sua bibliografia.

<sup>104</sup>Con "20T1" si intende "2020-Trimestre 1" e così di seguito. Successivamente, si ha, in modo del tutto analogo, "20S1" corrispondente a "2020-Settimana 1".

<sup>105</sup>Sulla definizione di ACF e PACF, e per la determinazione dell'ordine di un AR tramite questi si vedano le Appendici.

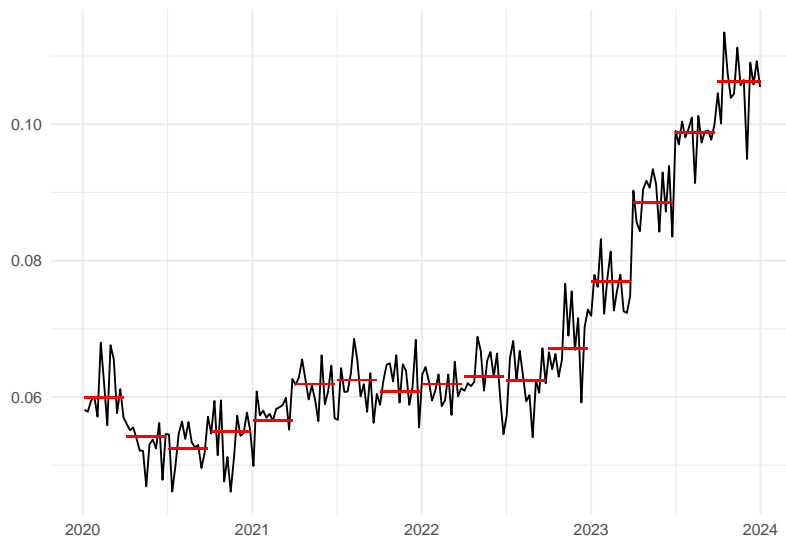


Figura 13: La serie ANIA settimanale, simulata a partire dai dati trimestrali secondo la procedura descritta.

Il campione simulato è di 209 valori anziché 208 - 52 settimane per 4 anni -, questo per uniformare temporalmente la serie alle variabili che verranno introdotte in seguito, le quali hanno 209 osservazioni per un numero corrispondente di settimane. La settimana aggiuntiva è stata data all'anno 2023, come ultimo valore, dato che è per quell'anno che le altre serie presentano 53 settimane.

Allo stesso modo si ottengono le altre due simulazioni, che per brevità qui non si riportano.

L'intenzione è ora quella di sfruttare il nuovo campione simulato per completare le tre serie osservate. Si deve definire un ulteriore procedimento e dare nuove ipotesi per raggiungere il risultato voluto.

**2) La Congiunzione delle serie:** Il passaggio, in questo caso, è ancora più delicato, poiché, a seconda di come vengono unite le serie, l'istante di transizione tra regimi si sposta. Questo provoca una distorsione nelle stime del modello per tutti i parametri. La scelta che è stata fatta è la seguente: la serie simulata e le tre HRG si suppone differiscano per un fattore additivo  $d$ , costante nel tempo ma specifico per i portafogli. L'idea alla base è che il processo simulato, grazie all'uso della varianza stimata, abbia lo stesso andamento del campione nelle date precedenti al 2023, ma che differisca per un effetto di traslazione. Data questa ipotesi, è necessario scegliere come stimare il fattore  $d$ . Volendo ridurre l'errore nel punto di congiunzione tra il campione e la serie simulata, tra il primo trimestre 2023 e l'ultimo trimestre 2022, lo stimatore che si è pensato di scegliere è la differenza tra  $x_{23T1}$  e l'indice di riscatto medio nel primo



trimestre 2023 dei tre portafogli  $\bar{y}_{23T1}$ <sup>106</sup>. In questo modo si possono far combaciare le due serie rispetto al primo trimestre 2023.

Per ottenere il valore di  $\bar{y}_{23T1}$  sarebbe necessario fare una media ponderata dei primi tredici tassi, usando come pesi le riserve medie delle diverse settimane - ricordando la definizione dell'indice di riscatto -, ma non si dispone di questo valore. Di conseguenza si è dovuto optare per una media aritmetica semplice.

Formalizzata, la procedura è la seguente<sup>107</sup>:

$$y_t^* = \begin{cases} x_t^* + d & \text{per } t = 20S1, \dots, 23S1 \\ y_t & \text{per } t = 23S2, \dots, 23S39 \end{cases},$$

$$d = \bar{y}_{23T1} - x_{23T1},$$

$$\bar{y}_{23T1} = \frac{1}{12} \sum_{s=23S2, \dots, 23S13} y_s,$$

dove  $y_t$  corrisponde a uno dei tre tassi delle HRG e  $y_t^*$  invece alla nuova serie simulata per la stessa HRG. Come si può notare, da  $y_t$  si è rimossa la prima settimana, e così anche per  $\bar{y}_{23T1}$ , ciò è stato fatto seguendo quanto detto in precedenza sulla prima settimana dell'anno del campione.

Seguendo il metodo descritto le tre nuove serie risultano complete fino al 2020. Vi è però un problema, visibile in particolare nella terza HRG completa, riportata in figura 14: si può notare una rottura nel punto di transizione tra la serie simulata e il campione originale; le due serie non combaciano in modo naturale.

Ciò è causato dal fatto che i primi valori delle tre HRG sono crescenti e conseguentemente la media è più lontana dai primi valori della serie. Pertanto, andando a calcolare la nuova serie, l'inizio della coda - cioè i valori più prossimi da sinistra al 2023 - sarà più in alto rispetto all'inizio del campione originale. Un approccio risolutivo potrebbe essere quello di utilizzare una media ponderata con pesi che diano più importanza ai primi valori del 2023, ma la loro scelta introdurrebbe un'ulteriore componente soggettiva. Si potrebbe anche pensare di utilizzare una media mobile per lisciare la serie, ma anche questo introdurrebbe nuova soggettività, oltre a ridurre la variabilità in punti non necessari. Si è scelto un criterio grafico, più semplificato: osservare come varia la media al variare del numero di settimane lasciate nel computo.

Guardando la figura 15, approssimativamente, per tutte e tre le serie si ha un *plateau* dal quinto al nono valore, quindi una scelta ragionevole è eliminare le ultime quattro settimane dal computo - un mese sui tre del trimestre -. Le tre nuove serie presentano una miglior transizione tra i due campioni, come si può constatare dalla figura 16.

---

<sup>106</sup>Nella descrizione della procedura si sta utilizzando il termine “stimatore”, ma è doveroso rimarcare il fatto che non si sta parlando di uno “stimatore statistico”, dato che allo stimatore definito dalla procedura non è associata alcuna distribuzione di probabilità.

<sup>107</sup>Più correttamente, sarebbe necessario distinguere la formula in base a quale HRG si riferisce. In questo caso si è scelto di non apporre nessun indice a tal fine, per non appesantire la simbologia.

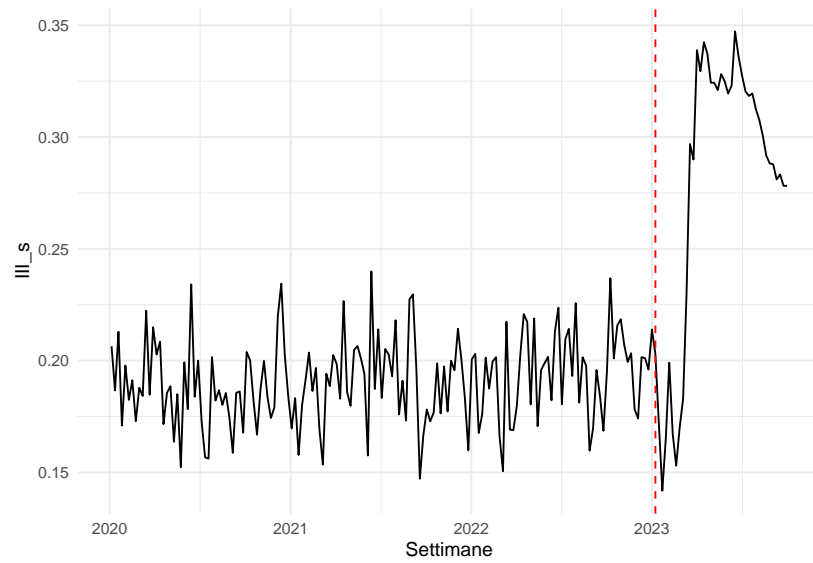


Figura 14: La terza HRG completata tramite la procedura.

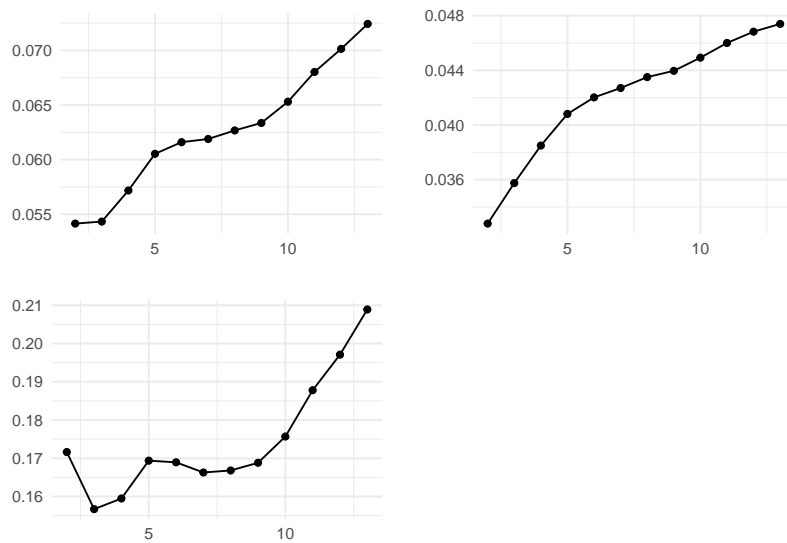


Figura 15: La media dei tassi di riscatto al variare del numero di settimane lasciate nel computo, per le tre HRG.

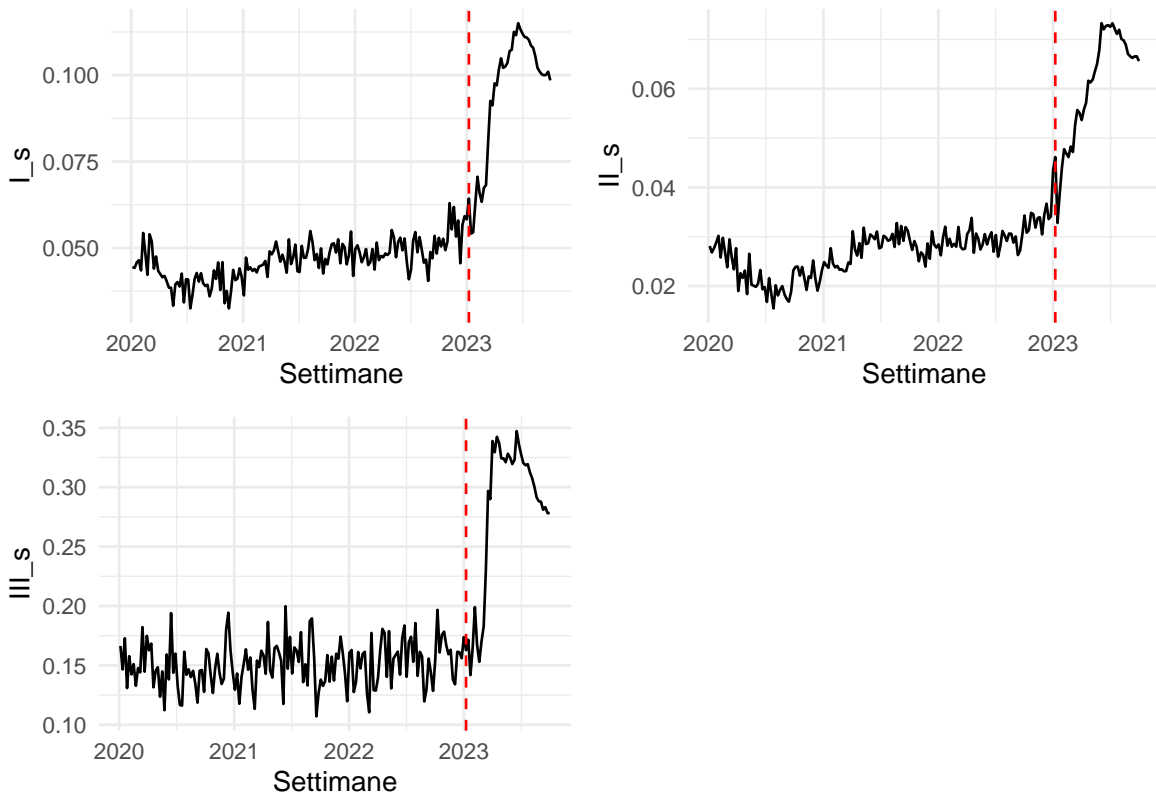


Figura 16: I grafici delle tre HRG aumentate grazie ai dati ANIA tramite la procedura descritta.

Questo è pertanto il campione definitivo. Si utilizzerà il termine di *pseudo-dati* e di *pseudo-campione* per rimarcare il fatto che, sebbene provengano da dati realistici, le nuove serie sono in parte soggette a distorsioni dovute all'approccio costruito e alle ipotesi fatte.

È bene però descrivere più nel dettaglio quali sono i punti critici della procedura e quali problematiche essa potrebbe portare all'analisi statistica. Una prima osservazione è che i due punti della procedura non tengono conto che la serie assume solo valori compresi tra zero e uno. Infatti, sia l'utilizzo della *Normale* che del fattore additivo  $d$  può portare, teoricamente, a valori negativi o sopra l'uno. La procedura potrebbe facilmente essere modificata per tenere conto di questo utilizzando una trasformazione delle variabili, ad esempio tramite "logit". Tuttavia, in questo caso, non si è ritenuto necessario, dato che il *range* di valori è sufficientemente ristretto da rendere improbabili valori errati; inoltre, le diverse trasformazioni avrebbero prodotto risultati indistinguibili da quelli ottenuti. Un secondo problema è legato alla disgregazione temporale: questa suppone di fatto che il processo sia indipendente, quando invece, trattandosi di una serie temporale, ci si potrebbe aspettare una forma di autocorrelazione. Le conseguenze di tale scelta potrebbero essere quelle di avere parametri non significativi stimati per il modello applicato. Altro problema ancora è la procedura di congiungimento delle due serie: la scelta provoca lo spostamento dell'istante di transizione e ciò può comportare la distorsione delle stime dei parametri o persino rendere i parametri non significativi. Un ultimo problema è legato alla stima del parametro  $d$  e della varianza: il primo dipende dalla media aritmetica delle prime settimane del campione ed è quindi legato ai problemi sopra discussi, la seconda è invece legata al modello base scelto; in particolare per la terza serie sembra causare una variabilità maggiore di quella che ci si potrebbe aspettare.

Tutti questi problemi sono stati tenuti conto nella valutazione delle stime dei modelli.

## 5.2 Le Variabili Esterne

Gli RSM richiedono di assegnare una variabile le cui variazioni causino le transizioni tra regimi. Qui sono state selezionate tre serie da testare:

1. Il rendimento settimanale dei BTP a dieci anni,
2. Il *Google Trend* delle ricerche su internet della parola "*Eurovita*",
3. L'indice dei riscatti stesso, preso con un certo ritardo.

La scelta delle variabili è stata indirizzata seguendo quanto descritto nel secondo capitolo, nell'ultimo paragrafo, dove sono stati riportati i possibili fattori che possono

causare l'evento di massa. La prima variabile è stata scelta con riferimento alla *Money-ness Hypothesis*, o all'ipotesi di tasso<sup>108</sup>. Ripetendo quanto scritto in precedenza, tra il 2022 e il 2023 si è avuto l'aumento da parte della Banca Centrale dei tassi di interesse, al fine di ridurre l'inflazione. La manovra ha reso più favorevole il riscatto rispetto al portare a termine la polizza e questa può essere una possibile causa - non l'unica - dello *shock* dell'indice di riscatto. Sarebbe più corretto l'utilizzo della differenza tra tasso di mercato e tasso accreditato, in questo caso però le polizze che costituiscono il portafoglio sono a tasso zero, perciò è sufficiente utilizzare un tasso di mercato. Il tasso che è stato scelto è la serie storica, dal 2020 al 2024, dei rendimenti dei BTP italiani con scadenza a 10 anni, con frequenza settimanale. Il grafico era già stato mostrato nel capitolo precedente, tuttavia lo si riporta nuovamente, in figura 17, per poter commentare al meglio l'evoluzione storica della serie.



Figura 17: La serie storica dei BTP a 10 anni, rendimenti settimanali, dal 2020/01/01 al 2024/04/07.

Si può constatare, come nel caso della serie dei riscatti, che il rendimento è cominciato a crescere nel 2022 raggiungendo un picco attorno a fine 2022; è rimasto poi stabile, sempre ad un valore alto, nel corso del 2023. I BTP però hanno raggiunto il loro secondo regime qualche settimana prima dei riscatti e ciò suggerisce che la serie debba essere presa con un certo differimento per essere usata come processo indipendente.

La seconda variabile scelta è invece legata al fenomeno del *panic selling* dovuto a eventi catastrofici<sup>109</sup>. In questo caso l'evento a cui ci si riferisce è il caso, riassunto nel-

<sup>108</sup>Si veda appunto il secondo capitolo, ultimo paragrafo.

<sup>109</sup>Si faccia ancora riferimento al secondo capitolo, ultimo paragrafo.

l'introduzione, del crollo di *Eurovita*, avvenuto indirettamente per effetto dell'incremento dei tassi, quindi in parte collegato alla variabile precedente. Anche qui richiamando quanto già detto, l'IVASS, precedentemente il crollo, aveva rimosso, a tutela degli assicurati, il diritto di riscatto per un gruppo di polizze; il congelamento del diritto e il panico generale prodotto dal successivo crollo dell'impresa hanno generato un clima di insicurezza verso le compagnie assicuratrici e in questo si può riconoscere un'ulteriore prodromo del riscatto in massa del 2023.

Per poter rappresentare tale fenomeno si è scelto di utilizzare l'“indice di interesse” prodotto da *Google Trend*: il sito costruisce un indice rappresentativo dell'interesse generale su un dato argomento, in base al numero di ricerche fatte sul motore di quel particolare argomento. La serie storica che ne risulta è normalizzata a valori reali tra zero e cento per essere interpretabile<sup>110</sup>. Per la parola “Eurovita”, il sito produce la serie riportata in figura 18.

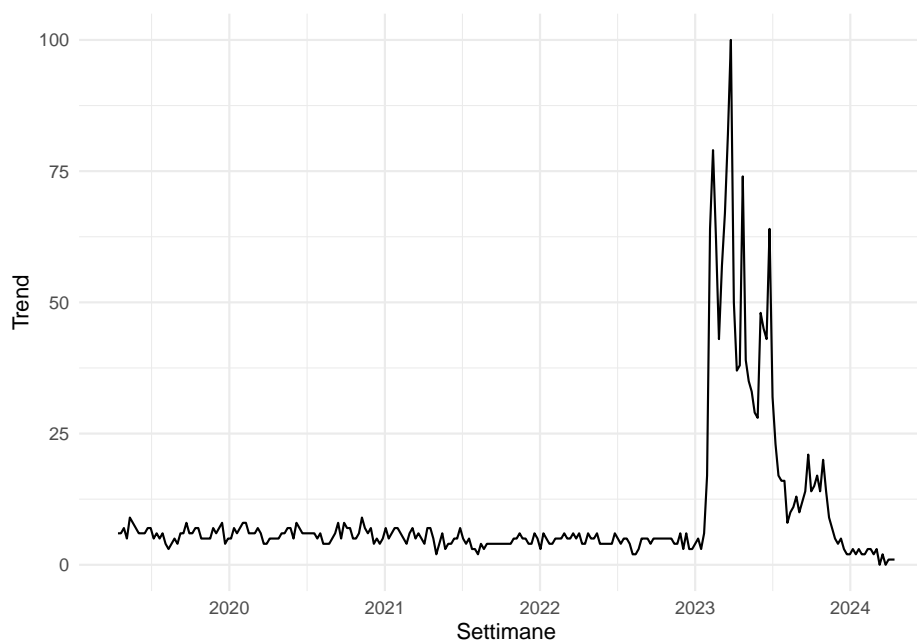


Figura 18: Il trend delle ricerche Google della parola Eurovita, 2019/04/14-2024/04/14.

Il grafico mostra come il maggior numero delle ricerche è stato fatto proprio nel corso del 2023, con qualche settimana di ritardo rispetto all'incremento in massa dei riscatti.

Il difetto di questa variabile è di non poter essere utilizzata, per ovvie ragioni, in un modello interno. Due candidati che potrebbero assurgere a tale ruolo sono: il *solvency Ratio* di mercato o il *solvency Ratio* della impresa stessa, l'ultimo disponibile agli assicurati, o ancora un indice di fiducia del settore assicurativo in generale.

<sup>110</sup>Sulla costruzione dell'indice e per altre possibili domande, si rimanda al sito: <https://trends.google.com/trends/>.

L'ultima scelta di variabile indipendente è l'indice di riscatto stesso preso con un certo ritardo. A tale scelta è legata l'ipotesi di effetto di correlazione e contagio in presenza di condizioni estreme. Il fenomeno in questione è legato con le due variabili precedenti ma si costituisce comunque come un fattore causale a sé stante. Ricordando la dinamica come presentata nel secondo capitolo: quando un assicurato osserva attorno a sé un incremento inatteso dei riscatti, anche senza conoscerne il motivo, aumenta la sua probabilità di riscattare la polizza. Per catturare il fenomeno si va ad utilizzare la serie stessa, e il ritardo indicherà la rapidità con cui si trasmettono le informazioni in merito e quanto velocemente in media gli operatori agiscono. Gli RSM che fanno utilizzo di tale variabile, si ricorda, sono i cosiddetti SETAR e SESTAR.

### 5.3 La Stima dei Modelli

Per cominciare viene fatta una verifica grafica dell'autocorrelazione, tramite ACF e PACF e grafico di dispersione per diversi ritardi<sup>111</sup>. Non verranno presentati tutti i grafici, ma solo quelli che si intende commentare.

I tre grafici delle diverse funzioni di autocorrelazione sono molto simili. Osservando in figura 19 quello riferito alla prima serie, si può notare che l'autocorrelazione è decrescente nel tempo, ma non esponenzialmente, come nel caso di un AR, ma quasi linearmente.

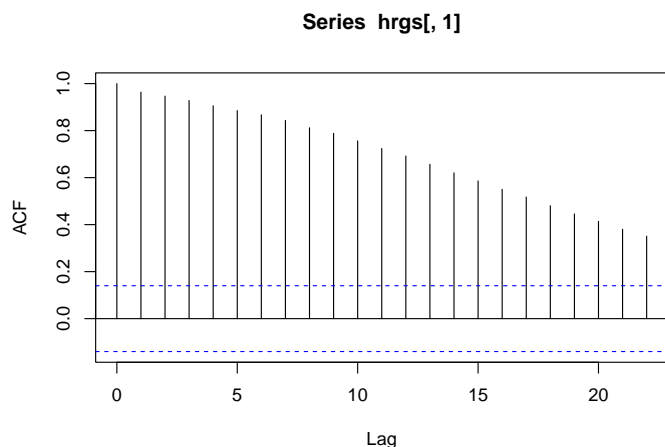


Figura 19: PACF della prima HRG.

<sup>111</sup>Per la definizione di ACF e PACF si rimanda alle Appendici.

Questo è solitamente sintomo di un processo non stazionario, come in effetti ci si aspettava. Ugualmente interessanti sono i grafici delle funzione di autocorrelazione parziale: sono presenti due picchi per tutte e tre le serie, ad indicare come possibile modello un AR di ordine 2<sup>112</sup>. In figura 20 si riporta la PACF del primo campione.

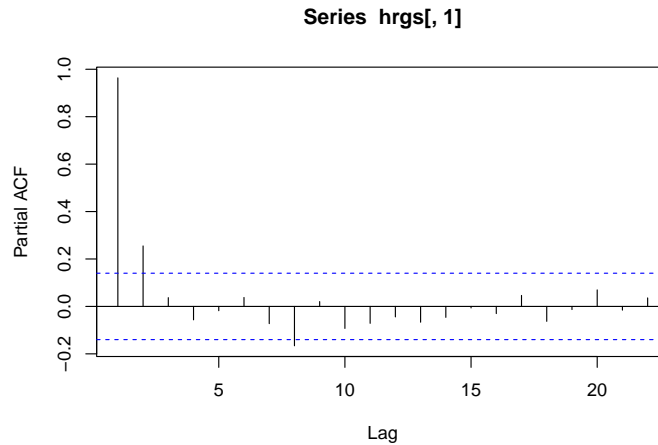


Figura 20: PACF della prima HRG.

A completare l'analisi si presentano in figura 21 anche i grafici a dispersione tra la variabile e un suo ritardo; per semplicità si riportano soltanto quelli riferiti alla prima HRG, per quattro ritardi.

I punti sono distribuiti in modo abbastanza lineare, se non per quell'insieme di punti intermedio che separa le due nuvole di punti: quel gruppo corrisponde proprio ai valori situati a transizione tra i due regimi, che rompono la linearità.

Dall'analisi delle autocorrelazioni ci si sposta all'analisi delle correlazioni tra le variabili che compongono il *dataset*. Si riporta in figura 22 il grafico delle correlazioni.

Tutte le variabili sono fortemente correlate positivamente. Ciò suggerisce che le due variabili esterne possono essere delle buone variabili esplicative per l'evento di massa.

Dopo aver svolto l'analisi esplorativa del campione, si può ora passare alla stima dei modelli applicati allo pseudo-campione, partendo dai modelli Autoregressivi, per verificare l'ipotesi di non linearità. I diversi approcci saranno svolti e commentati nello specifico per la prima serie, questo per rendere più scorrevole il discorso. Le diverse procedure sono state applicate allo stesso modo anche alle altre due HRG rimanenti, i loro risultati sono stati riportati però in modo sintetico.

### 5.3.1 La stima del modello AR

L'indice di riscatto richiede di essere trasformato per poter applicare correttamente i modelli visti, poiché essi richiedono valori reali. La trasformazione comunemente

<sup>112</sup>In merito alla determinazione dell'ordine di un AR si rimanda alle Appendici.



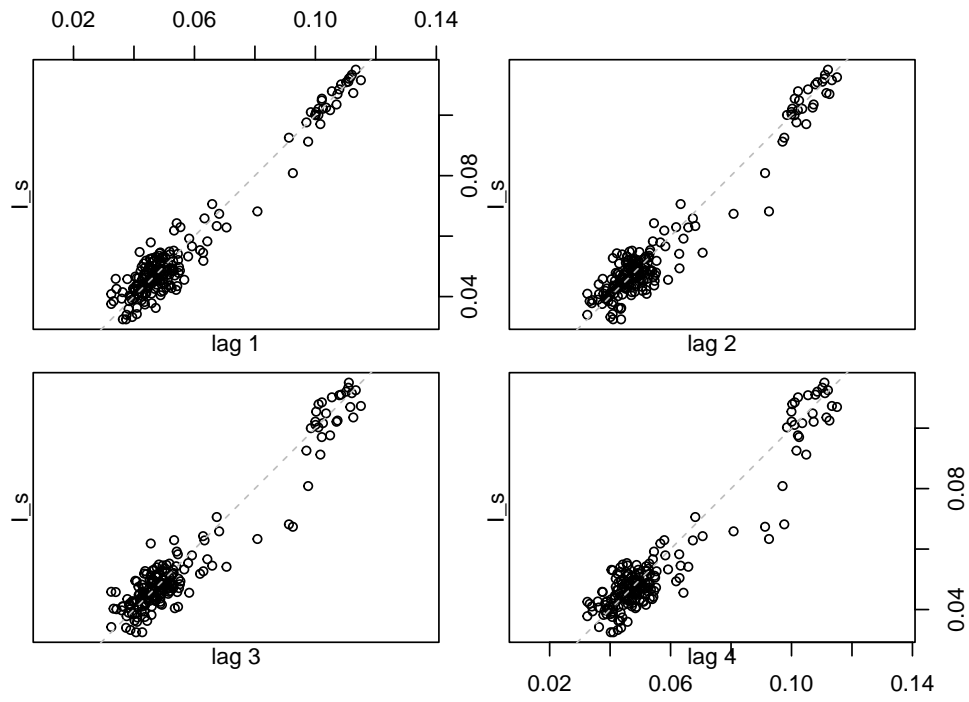


Figura 21: HRG I contro la serie stessa, ritardata.

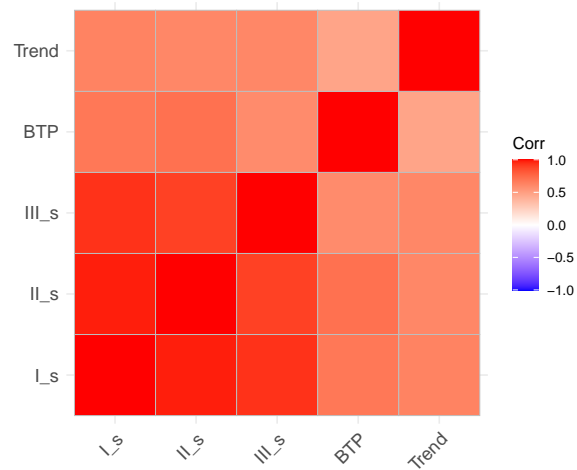


Figura 22: Grafico delle correlazioni tra variabili.

utilizzata in approcci simili è il “*logit*”, che è la funzione inversa della logistica standard ed è così definita<sup>113</sup>:

$$\text{logit}(x) = \ln\left(\frac{x}{1-x}\right) .$$

Il passaggio della variabile al *logit* rende più complessa l’interpretazione dell’impatto delle variabili esplicative sulla variabile dipendente e soprattutto rende più difficile ottenere i momenti di quest’ultima, e quindi valutare il *mass lapse*. Per risolvere questo problema si può ricorrere a diversi espedienti: l’utilizzo dei “*Generalized ARMA Models*” - l’equivalente dei GLM nell’ambito delle serie storiche - , la valutazione diretta o simulata della distribuzione trasformata oppure l’utilizzo del polinomio di *Taylor* per il calcolo approssimato dei momenti. Il primo metodo in particolare difficilmente può essere qui applicato poiché non esiste attualmente una implementazione dei modelli RSM con questi ultimi.

In realtà, come si vedrà meglio dalle stime dei modelli, non è necessario andare a trasformare la variabile perché il *range* in cui si muovono le tre HRG è relativamente piccolo, così come la loro variabilità nei regimi è piccola, e ciò rende altamente improbabili valori *out of bounds* per i modelli stimati. Questo è valido solo in linea approssimata e soprattutto non vale in generale per qualsiasi campione.

Applicando la trasformazione *logit* alla prima serie, si può vedere dal grafico in figura 23 che la trasformazione ottenuta è di fatto lineare, poiché si modificano solamente la posizione e la scala della curva osservata.

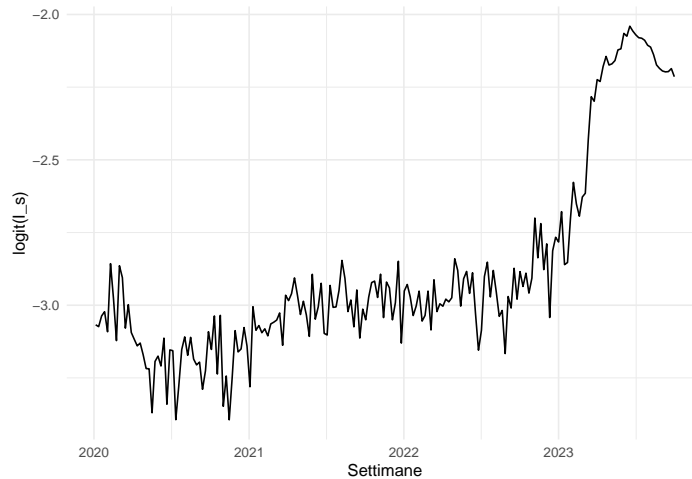


Figura 23: Trasformazione tramite funzione logit della prima HRG.

Data la variazione molto piccola della variabile, la funzione può essere approssimata abbastanza fedelmente da un polinomio di grado uno. Ciò fa sì che le stime dei parametri autoregressivi risultino più o meno equivalenti alle loro controparti nel modello non trasformato.

<sup>113</sup>Non è l’unica trasformazione possibile: come nel caso dei GLM, si può utilizzare ad esempio la FdR di una *Normale* o una qualsiasi FdR continua.

Se l'andamento è approssimativamente uguale, ci si può aspettare che i modelli presentino più o meno le stesse stime per i parametri autoregressivi e, nel caso dei RGM, condividano l'istante a cui è legato il cambio di regime.

Detto ciò, stimando un modello AR di ordine due, per la serie trasformata al *logit* e non, si ottengono i seguenti risultati:

Tabella 1: AR(2) di I\_s

	Estimate	Std. Error	t value	Pr(> t )
<b>ar1</b>	0.5883	0.06531	9.008	0
<b>ar2</b>	0.4104	0.06604	6.214	5.18e-10
<b>intercept</b>	0.0004723	0.000893	0.5288	0.5969

Tabella 2: AR(2) di logit(I\_s)

	Estimate	Std. Error	t value	Pr(> t )
<b>ar1</b>	0.5276	0.06359	8.296	0
<b>ar2</b>	0.4601	0.06427	7.159	8.118e-13
<b>intercept</b>	-0.0291	0.05858	-0.4967	0.6194

L'intercetta è non significativa per entrambi i modelli, mentre i due parametri autoregressivi sono significativi e molto simili nonostante la trasformazione, ciò appunto per quanto detto poco sopra. I due modelli sono però non stazionari, come ci si poteva aspettare. Questo lo si può verificare utilizzando il test di *Dickey-Fuller* aumentato<sup>114</sup>:

Tabella 3: Augmented Dickey-Fuller Test per I\_s

Test statistic	Lag order	P value	Alternative hypothesis
-1.373	2	0.8385	stationary

Tabella 4: Augmented Dickey-Fuller Test per logit(I\_s)

Test statistic	Lag order	P value	Alternative hypothesis
-1.848	2	0.6399	stationary

Si rifiuta per entrambe le serie l'ipotesi alternativa di stazionarietà. Anche osservando in figura 24 i residui dei due modelli si può vedere il punto di rottura con l'ultima parte del campione.

<sup>114</sup>Il test è stato citato nel capitolo precedente. Per una sua definizione si rimanda a Stock and Watson (2020).

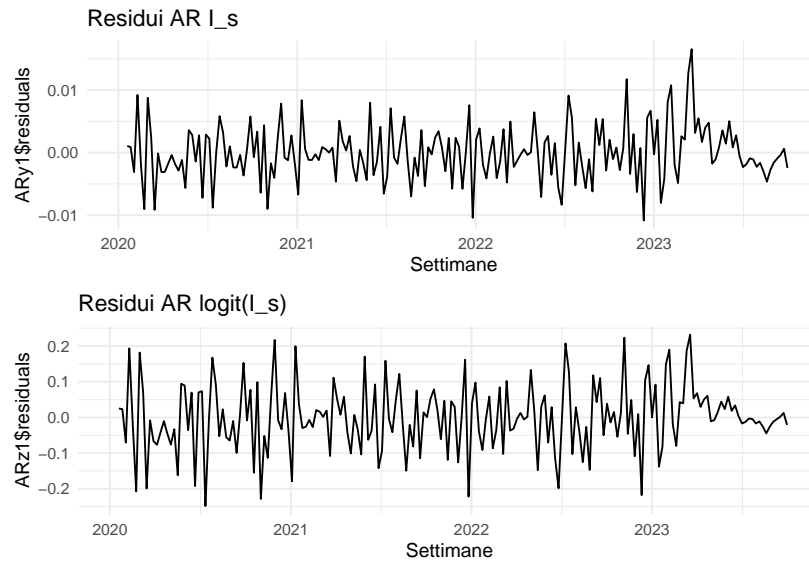


Figura 24: Grafico dei residui del modello AR semplice.

### 5.3.2 La stima del modello TAR

Spostandosi ora ai RGM, verranno applicati per primi i modelli a barriera, sempre alla prima HRG, utilizzando via via le tre possibili variabili esplicative. Per operare la stima verrà sfruttato il pacchetto R “NTS”, “*Nonlinear Time Series Analysis*”, in particolare la funzione “uTAR”, la quale permette di stimare il modello per diversi ordini dei regimi e diverse varianze per i regimi.

**Stima con variabile esterna BTP:** Utilizzando come processo indipendente i BTP a ritardo uno, per due regimi e ordine dei regimi pari a due, si riportano le stime per i modelli con - tabella 5 - e senza trasformazione - tabella 6 - .

Tabella 5: TAR per I\_s, 2 regimi, variabile BTP a un ritardo

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.01356	0.003242	4.183	5.087e-05
<b>b1</b>	0.3348	0.07894	4.242	4.045e-05
<b>b2</b>	0.366	0.07975	4.59	9.884e-06
<b>a0</b>	0.005438	0.002202	2.47	0.01696
<b>a1</b>	0.7381	0.1255	5.879	3.376e-07
<b>a2</b>	0.214	0.1232	1.737	0.08851

Tabella 6: TAR per  $\logit(I_s)$ , 2 regimi, variabile BTP a un ritardo

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	-0.869	0.2147	-4.047	8.599e-05
<b>b1</b>	0.3443	0.07919	4.347	2.656e-05
<b>b2</b>	0.3711	0.07985	4.648	7.753e-06
<b>a0</b>	-0.133	0.07141	-1.863	0.06835
<b>a1</b>	0.6199	0.1143	5.425	1.69e-06
<b>a2</b>	0.3154	0.111	2.842	0.006472

Con **b** sono stati indicati i parametri del regime basso, mentre con **a** quelli del regime alto; inoltre, con zero è indicata l'intercetta mentre con 1 e 2 rispettivamente il coefficiente per il ritardo 1 e 2. La barriera è la medesima per entrambi i modelli: 0.03851, così come sono approssimativamente gli stessi anche i parametri autoregressivi e il numero di unità per regime: 141 per il regime basso e 53 per il regime alto; le varianze e le intercette invece sono diverse. Una criticità che si può notare è che il secondo regime di entrambi i modelli ha **a0** e **a2** non significativi per un livello del 1%. Questo potrebbe suggerire che un ordine 1 per il regime alto potrebbe essere più adeguato.

Anche qui si possono studiare i residui, riportati in figura 25: per entrambi i modelli l'ACF e la PACF confermano l'incorrelazione di questi. In aggiunta, il grafico della serie temporale dei residui non presenta più la rottura che si osservava per il modello AR.

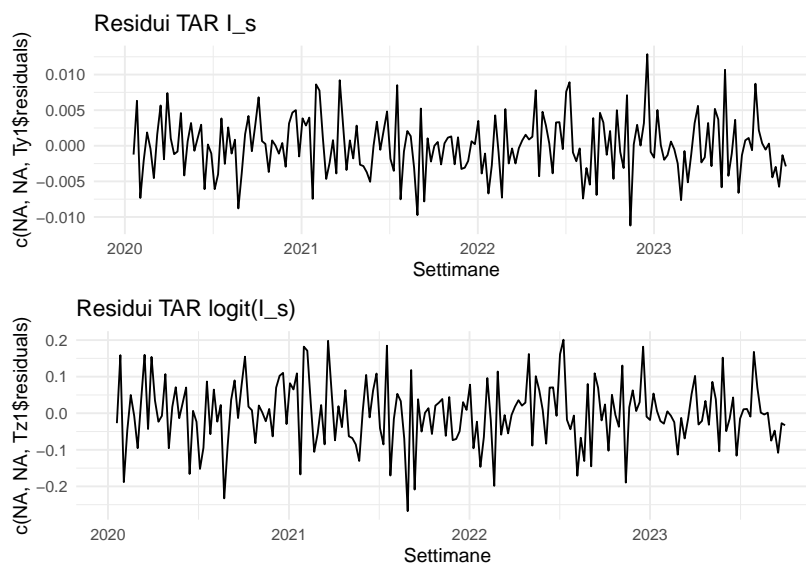


Figura 25: Grafico dei residui del modello TAR stimato con e senza trasformazione logit.

Un buon modo per valutare se il modello risulta adeguato è quello di simulare il processo utilizzando i parametri stimati. Si ipotizza una distribuzione *Normale* per i residui. Per avere la confrontabilità tra i modelli, verrà invertita la trasformazione applicata al secondo modello. I grafici sono riportati in figura 26.

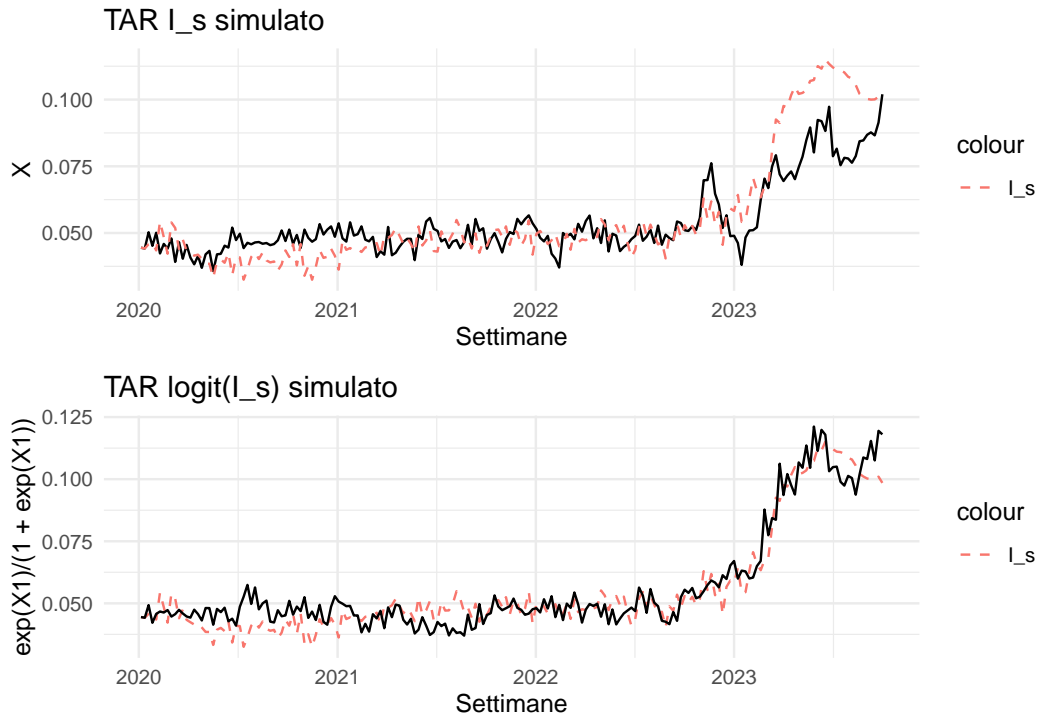


Figura 26: Simulazioni dal modello TAR stimato, con e senza trasformazione.

Entrambi i modelli sembrano riprodurre l'andamento del campione, anche la simulazione per la variabile non trasformata, seppur in modo peggiore. Si può notare, tuttavia, che entrambe non colgono perfettamente l'inizio della serie, questo perché il suo andamento è probabilmente legato ad un'altro fattore. Inoltre, il secondo regime del primo modello sembra essere molto variabile, forse non stazionario, questo è difficile da valutare poiché il secondo regime ha solo 53 valori. Come detto, non è necessaria, per la stazionarietà di un TAR, la stazionarietà di entrambi i regimi, è però comunque importante tenerne conto, dato che il secondo regime è legato alla stima dell'evento estremo. Senza aggiungere ulteriori tabelle, si può vedere che usare un ordine 1 per il secondo regime riduce questo problema e aumenta la significatività dei parametri.

Il differimento del processo indipendente è stato preso pari a uno per comodità ma sarebbe necessaria una verifica migliore della scelta. Non esistendo una procedura di stima statistica del parametro  $d$ , un metodo semplice è la verifica grafica tramite confronto delle due serie; perché il confronto sia sensato, le variabili sono riscalate per avere le stesse dimensioni. Il grafico di confronto è riportato in figura 27.

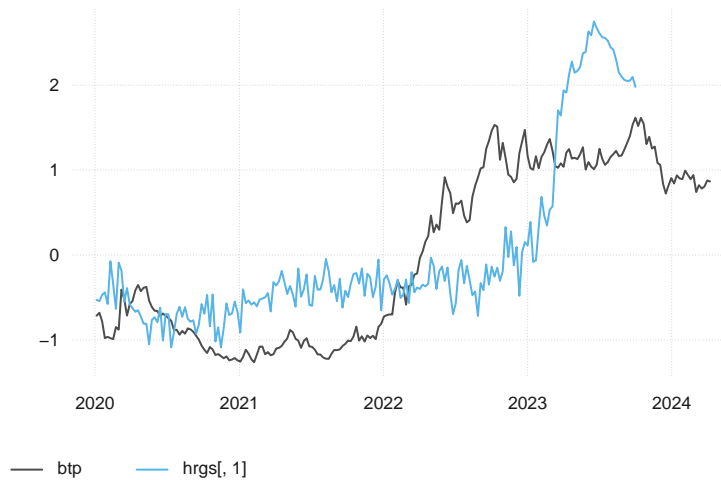


Figura 27: Grafico riscalato dell'indice di riscatto e dei rendimenti BTP.

Ragionando in termini approssimativi, il grafico potrebbe suggerire un differimento maggiore, tenuto conto anche del fatto che la componente autoregressiva dei TAR rende la transizione non immediata. Variando il numero di settimane di differimento, si è ottenuto come buon risultato, con ordine 1 nel secondo regime, un valore di  $d = 13$ , circa un trimestre, che alza la soglia a 0.04036, e abbassa leggermente la stima del parametro **a1**. Per altri differimenti superiori la soglia si abbassava a livelli poco realistici, ma in generale il modello rimane con un buon fitting.

Un criterio che qui si è utilizzato per valutare la bontà del modello, nell'analisi di questo specifico pseudo-campione, è osservare quanto è alta la stima della soglia rispetto all'andamento medio: se infatti si ottenesse un valore contenuto, ciò avrebbe come effetto sulla serie dei riscatti di osservare spesso le transizioni tra regimi, quando invece, da quanto osservato in storicamente, l'evento è poco frequente. Anche per queste tre HRG, preso ad esempio il BTP come processo indipendente, se la soglia si abbassa allora l'istante di transizione si allontana dal 2023, verso le date dove la serie era più stazionaria e dove sembra non esserci transizione.

Anche il numero dei regimi è stato scelto pari a due ma sarebbe più corretto farne una verifica. Per rendere più scorrevole il discorso, non si riporteranno tabelle, risulta però che per questo TAR - ma anche per i modelli futuri - il terzo regime aggiuntivo è sempre stimato non significativo.

Dato che il livello di *fitting* ottenuto dai due modelli - con e senza trasformazione - è simile, si procederà nei successivi punti presentando per semplicità il solo modello non trasformato. Le possibili differenze riscontrabili saranno riportate, mentre, nella valutazione dell'impatto del rischio di estinzione di massa, verranno utilizzati tutti e due i modelli.

**Stima con variabile esterna il Google Trend di Eurovita:** Passando alla seconda variabile, questa, rispetto alla precedente, sembra presentare il suo incremento ad inizio 2023, un po' dopo rispetto ai tassi di riscatto, come si vede dal grafico di confronto riportato in figura 28.

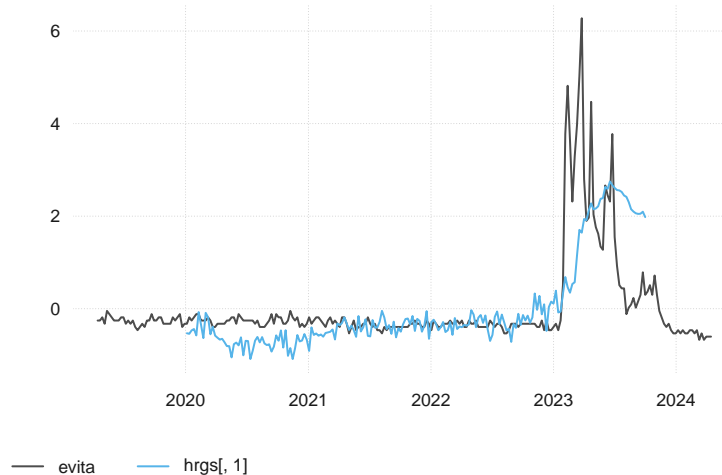


Figura 28: Grafico di confronto tra l'indice di riscatto e il Google trend di Eurovita.

Ciò può avere più spiegazioni: o la variabile non coglie il reale inizio dell'interesse verso la situazione - cosa difficile da sostenere, dato l'elevato utilizzo attuale di internet -, o i riscatti erano già crescenti in precedenza - per effetto dei tassi di mercato ad esempio - mentre il picco estremo effettivo è stato causato dal crollo di *Eurovita*, quindi un'azione congiunta dei due processi, oppure semplicemente la variabile è solo correlata col fenomeno e non rappresenta la reale causa del cambio di regime dei riscatti.

Per verificare quanto detto si può stimare il modello TAR con la variabile esterna il *Google Trend*, per ritardo 1<sup>115</sup>; le stime sono riportate in tabella 7. Si potrebbe scegliere anche nessun ritardo, ma questo presupporrebbe la contemporaneità degli eventi e ciò non sarebbe interessante ai fini dell'analisi del rischio di riscatto.

<sup>115</sup>La stima di un TAR di per sé non permette la verifica dell'effetto causale, ma solo di una possibile correlazione. Per poter misurare l'effetto causale sarebbero necessarie ulteriori ipotesi, poter modificare la struttura del TAR per accogliere più regressori e svolgere dei rigorosi test. Attualmente questo non è possibile tramite la funzione di R utilizzata, pertanto qui ci si limita ad una valutazione della correlazione in termini temporali tra i due fenomeni. Di ciò si è già discusso nel precedente capitolo, nel paragrafo sulle criticità degli RSM.



Tabella 7: TAR per I\_s, 2 regimi, variabile Eurovita a un ritardo

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.005314	0.001878	2.829	0.00527
<b>b1</b>	0.427	0.07075	6.036	1.085e-08
<b>b2</b>	0.4627	0.06987	6.622	5.245e-10
<b>a0</b>	0.01053	0.00397	2.653	0.01248
<b>a1</b>	0.9042	0.04044	22.36	1.1e-20

Il valore stimato per la barriera è 10, coerente con l'istante in cui la variabile Eurovita comincia il picco. Viene superato per la prima volta nella data 2023-01-29, mentre, nel TAR con BTP, la data era 2022-09-18, molto prima quindi.

I parametri sono tutti significativi a livello dello 5% e, facendo una rapida verifica grafica dei residui, questi appaiono incorrelati e uniformemente dispersi. Più interessante è la simulazione, riportata in figura 29.

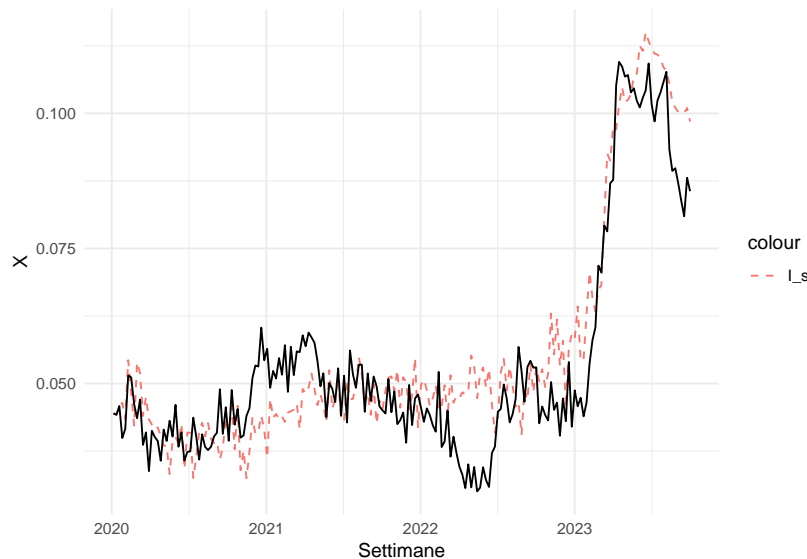


Figura 29: Simulazione del modello TAR stimato, con variabile esterna Eurovita.

Anche in questo caso il processo simulato sembra riprodurre in modo simile l'andamento del campione, ma con una maggior volatilità nella parte iniziale, riferita al primo regime. Ciò è dovuto al fatto che l'AR del primo regime ha un andamento quasi non stazionario. Modificando il differimento le stime peggiorano, perciò si è ritenuto come modello migliore questo mostrato.

Confrontandolo con il modello precedente, con variabile esterna i BTP, sembra essere peggiore, in termini di stima dei parametri.

**Stima del modello SETAR:** Passando al modello SETAR, la scelta del ritardo da applicare diventa ancora più critica. Si è perciò ritenuto necessario utilizzare una procedura più strutturata ed un criterio più oggettivo. Tramite il pacchetto R “tsDyn” è possibile automatizzare la procedura di selezione di  $d$  tramite la funzione “*selectSETAR*”. Il criterio utilizzato si basa sul confronto dell’AIC prodotti dal modello per diversi ritardi. La procedura richiede di fissare un range per  $d$  e in questo caso, visto quanto appreso per il TAR con variabile esterna BTP, gli estremi scelti sono una settimana ( $d = 1$ ) a un trimestre ( $d = 13$ ).

Tabella 8: Selezione del differimento

d	p_1	p_2	B	AIC pooled
2	2	1	0.06296	-1568
3	2	1	0.05792	-1568
3	2	1	0.05917	-1566
1	2	1	0.06427	-1565
1	2	1	0.06286	-1565
1	2	1	0.06587	-1564
1	2	1	0.05827	-1564
3	2	1	0.06184	-1564
3	2	1	0.06296	-1564
1	2	1	0.06296	-1564

La tabella 8 fornisce come possibili candidati per il ritardo da utilizzare i valori 2 e 3<sup>116</sup>. Avendo stimato i due modelli per i due ritardi, si è ritenuto migliore il secondo, poiché i parametri stimati risultano significativi per un livello di confidenza maggiore e il secondo regime ha il parametro autoregressivo più piccolo, a indicare un regime più stazionario. Di seguito la tabella 9 con le sue stime:

Tabella 9: SETAR per I\_s, 2 regimi, a ritardo tre

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.007836	0.002756	2.844	0.005052
<b>b1</b>	0.4357	0.07234	6.023	1.173e-08
<b>b2</b>	0.3995	0.07433	5.374	2.734e-07
<b>a0</b>	0.01782	0.003536	5.039	1.921e-05
<b>a1</b>	0.8332	0.03534	23.58	2.324e-21

<sup>116</sup>La procedura fornisce anche per lo stesso ritardo, diverse stime possibili per la barriera B.

Anche i residui del modello sembrano essere molto buoni. Si riporta in figura 30 la simulazione del processo.

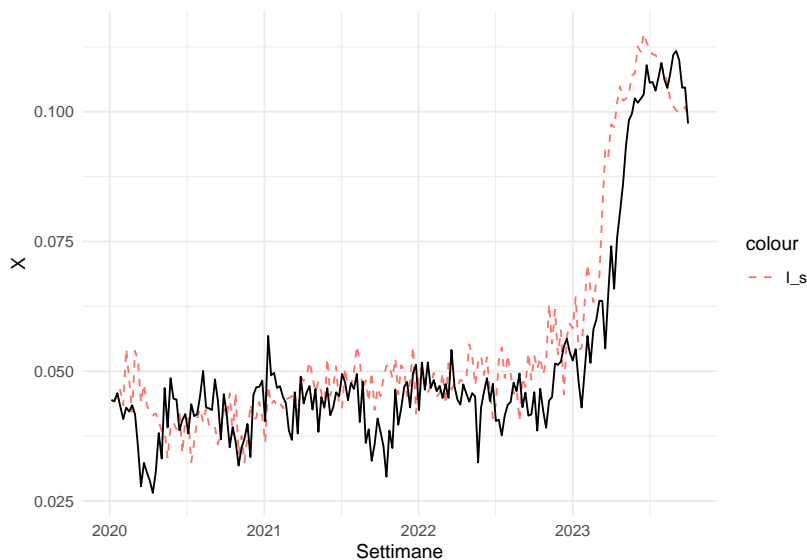


Figura 30: Simulazione del modello SETAR stimato.

La serie simulata sembra rappresentare bene lo pseudo-campione. Bisogna però osservare che l'istante di transizione tra regimi per i SETAR è legato alle fluttuazioni della variabile stessa e non da fattori esterni e ciò rende forse troppo prevedibili - cosa poco realistica - gli istanti in cui avverrà la transizione; il passaggio diventa un fenomeno ciclico, con frequenza costante, in termini attesi. Ugualmente, le previsioni sull'andamento futuro del SETAR, dipendono da come fluttua il secondo regime, e non tanto da azioni esterne, come interventi della compagnia assicurativa.

A questo proposito si propone in figura 31 una seconda simulazione nella quale invece non si è verificato il salto, poiché la variabile non ha raggiunto la soglia.

Qui il *fitting* è peggiore per la parte finale appunto perché la serie non ha superato la soglia e non si è verificata la transizione.

**Stima per la seconda e la terza HRG:** Le stesse procedure sono state seguite per la stima dei modelli TAR per le rimanenti due HRG. Sono state osservate delle sostanziali differenze nelle stime ottenuti e sono state necessarie delle variazioni al campione e ai modelli. Nello specifico, per la seconda HRG è risultato necessario rimuovere le prime 70 settimane, questo perché il processo presenta un particolare *trend* nel primo anno, visibile anche nelle altre due serie ma qui più pronunciato, che riduceva la significatività delle stime. Si è scelta la rimozione di tale segmento poiché l'interesse della Tesi è quello di valutare nello specifico la transizione tra regimi e di stimare il regime alto, non tanto di valutare l'andamento dei tassi nel regime basso. Sarebbe stato necessario inserire dei regressori che potessero dar conto di quelle osservazioni,

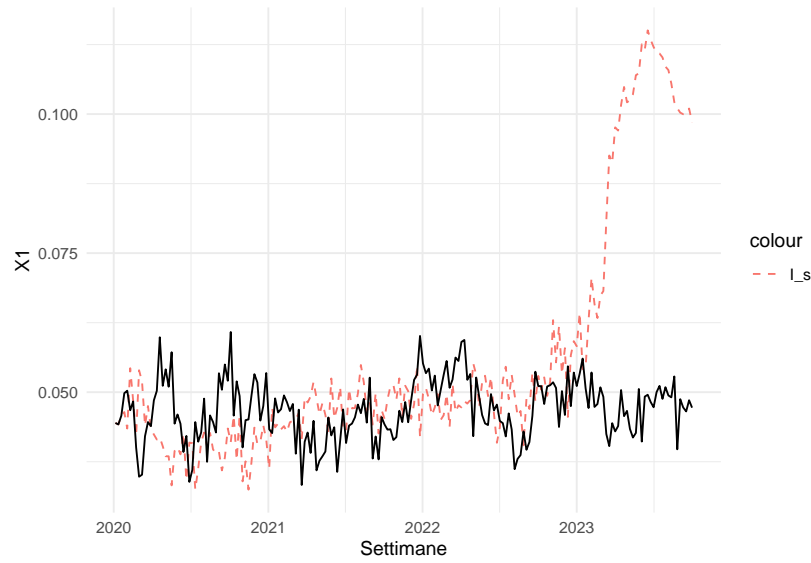


Figura 31: Una seconda simulazione del modello SETAR stimato.

ma, come detto, l'interesse è di valutare nello specifico il *mass lapse*. Oltre a questo, sempre per la seconda serie, il numero di parametri autoregressivi è stato posto pari a 1, mentre i differimenti ritenuti migliori per le tre variabili sono: 6 per BTP e 2 per il modello SETAR. La variabile Eurovita è stata rimossa poiché i modelli stimati sono stati ritenuti non adeguati. Nello specifico, l'incremento dei tassi precede di molto l'incremento del *Google trend* e ciò ha prodotto stime molto basse della barriera, sino a valori - tra il 6 e il 7 - giudicati non corretti rispetto all'effettivo andamento delle osservazioni.

Anche per la terza HRG sono stati necessari particolari aggiustamenti: la serie è stata mantenuta intera, mentre si è scelto anche qui un ordine 1 per entrambi i regimi. Il parametro autoregressivo per il primo regime è stato stimato molto piccolo e poco significativo. Le cause di questo sono state attribuite al processo di costruzione dello pseudo-campione ed in particolare al problema già discusso della forte variabilità osservata per la simulazione della terza HRG e per l'utilizzo di errori indipendenti. I ritardi per le variabili esplicative sono stati valutati come segue: 1 per BTP e 1 per i SETAR. Anche qui è stata rimossa la variabile Eurovita per le medesime ragioni della seconda serie.

### 5.3.3 Stima dei Modelli LSTAR

Per la stima dei modelli STAR si passerà ad utilizzare il pacchetto R già citato "tsDyn", che risulta essere l'unico disponibile, attualmente, su tale *software*. Il pacchetto implementa solamente la funzione logistica come funzione di transizione. Di seguito l'applicazione per le tre variabili.

**Stima con variabile esterna i BTP:** La stima del modello LSTAR, riportata in tabella 10, con variabile esterna BTP a ritardo 1 mostra già alcune problematiche che saranno comuni a tutti i modelli successivi provati degli STAR.

Tabella 10: LSTAR per I\_s, 2 regimi, variabile BTP, ritardo 1

	Estimate	Std. Error	t value	Pr(> z )
<b>b0</b>	0.01355	0.003861	3.51	0.0004476
<b>b1</b>	0.3789	0.08667	4.372	1.233e-05
<b>b2</b>	0.3206	0.06786	4.725	2.305e-06
<b>a0</b>	-0.01176	0.00604	-1.947	0.05149
<b>a1</b>	0.313	0.1294	2.42	0.01552
<b>g</b>	300	383.9	0.7814	0.4346
<b>B</b>	0.03529	0.003502	10.08	0

La stima del parametro  $\gamma$  è pari a 300, un valore molto alto. Come osservato nel capitolo precedente, quando tale parametro è elevato - tendente a più infinito - il modello STAR diventa indistinguibile dai modelli TAR - per la stessa variabile esterna - . Ciò potrebbe suggerire che quest'approccio potrebbe non essere adeguato per rappresentare il modello e che sarebbe da preferire il modello TAR.

La stima del parametro stesso è non significativa e, inoltre, la costante del secondo regime è negativa, oltre che non significativa, un valore non possibile per l'indice. La soglia risulta pari a 0.0352866, leggermente sotto quelle stimate dai precedenti modelli. Per i modelli STAR, una soglia bassa genera una distorsione maggiore nella stazionarietà del regime basso, a causa proprio della funzione di transizione  $G$  che definisce l'approccio. La ragione di tali stime singolari potrebbe risiedere nel fatto che il modello non sia in grado di rappresentare i dati.

Il modello è perciò stato valutato in generale non accettabile ed è da preferirsi l'equivalente nei modelli TAR. Si può vedere dalla tabella 11 che anche aumentando il differimento a 13 settimane per la variabile esterna, non si ottengono buoni risultati.

Tabella 11: LSTAR per I\_s, 2 regimi, variabile BTP, ritardo 13

	Estimate	Std. Error	t value	Pr(> z )
<b>b0</b>	0.01189	0.003621	3.283	0.001026
<b>b1</b>	0.4009	0.08494	4.72	2.36e-06
<b>b2</b>	0.3364	0.06972	4.825	1.402e-06
<b>a0</b>	-0.007319	0.00486	-1.506	0.1321

	Estimate	Std. Error	t value	Pr(> z )
<b>a1</b>	0.2373	0.09649	2.459	0.01392
<b>g</b>	277	206.6	1.341	0.1799
<b>B</b>	0.03284	0.003333	9.852	0

Tali problemi si possono meglio osservare simulando dal modello stimato. Dalla figura 32, la quale riporta una simulazione, si può vedere come la stima della costante del secondo regime sfalsa la simulazione; il modello è da rigettare.

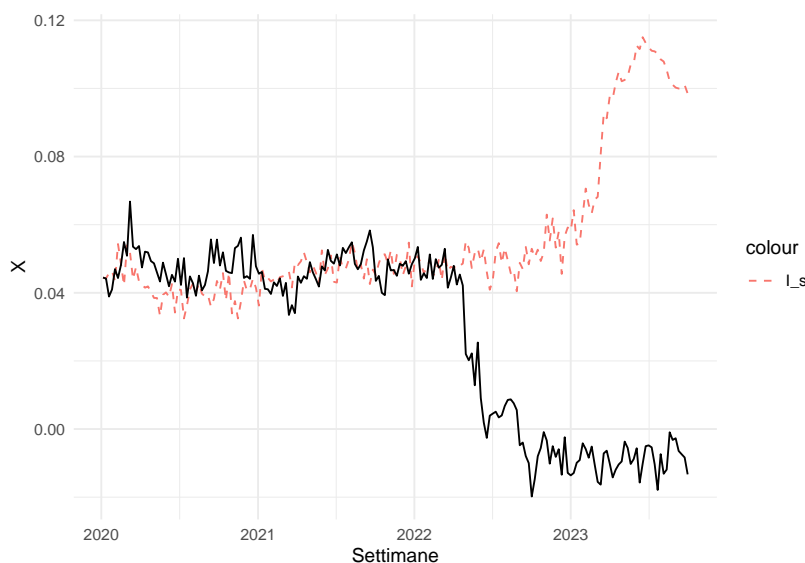


Figura 32: Simulazione del modello LSTAR stimato, con variabile esterna BTP.

**Stima con variabile esterna il Trend Google:** Anche in questo caso i risultati non sono buoni, ma migliori rispetto al caso precedente. Per il modello a ritardo uno, le stime vengono presentate in tabella 12.

Tabella 12: LSTAR per I\_s, 2 regimi, variabile EVITA, ritardo 1

	Estimate	Std. Error	t value	Pr(> z )
<b>b0</b>	0.006413	0.002208	2.904	0.003686
<b>b1</b>	0.4955	0.06832	7.253	4.086e-13
<b>b2</b>	0.3719	0.06644	5.598	2.174e-08
<b>a0</b>	0.004756	0.005454	0.872	0.3832
<b>a1</b>	0.03469	0.06862	0.5055	0.6132

	Estimate	Std. Error	t value	Pr(> z )
<b>g</b>	100	1695	0.059	0.953
<b>B</b>	9.991	0.1471	67.92	0

Il parametro  $\gamma$  è sempre non significativo, ma più basso del precedente; la stima della barriera è coerente con i risultati precedenti e tutti i parametri sono positivi, ma quelli del secondo regime sono non significativi. La criticità più grave risiede nella stima della costante del secondo regime: essa infatti, seppur positiva, è inferiore alla costante del regime basso e ciò genera gli stessi problemi riscontrati con il modello precedente. Il grafico dei residui, presentato in figura 33, presenta delle forti irregolarità verso la fine del campione

Seppur migliore del precedente, il modello è stato anch'esso rigettato, avendo anche osservato una sua traiettoria simulata.

**Stima del modello SESTAR:** Passando all'ultima variabile esplicativa, i risultati migliorano di molto, tuttavia i parametri del secondo regime sono non significativi.

Le stime del modello stimato con ritardo 3, in coerenza con il SETAR, sono riportate in tabella 13.

Tabella 13: SESTAR per I\_s, 2 regimi, ritardo 3

	Estimate	Std. Error	t value	Pr(> z )
<b>b0</b>	0.0116	0.003215	3.609	0.0003075
<b>b1</b>	0.4558	0.07342	6.208	5.373e-10
<b>b2</b>	0.2875	0.06889	4.174	2.996e-05
<b>a0</b>	0.008756	0.007914	1.106	0.2686
<b>a1</b>	0.06903	0.1034	0.6674	0.5045
<b>g</b>	300	87.22	3.44	0.0005823
<b>B</b>	0.06025	0.001841	32.73	0

Rimangono valide anche qui le considerazioni fatte per il modello SETAR, in merito alla stima della frequenza del *mass lapse*. Il modello è stato anch'esso rigettato.

**Stima per la seconda e la terza HRG:** Anche in questo caso si è svolta la verifica degli STAR sulle due serie rimanenti e le conclusioni ricavate sono le medesime fatte per la prima HRG, ovvero che il modello LSTAR non si adatta bene allo pseudo-campione e in particolare, con le sue stime, suggerisce di utilizzare i modelli TAR.

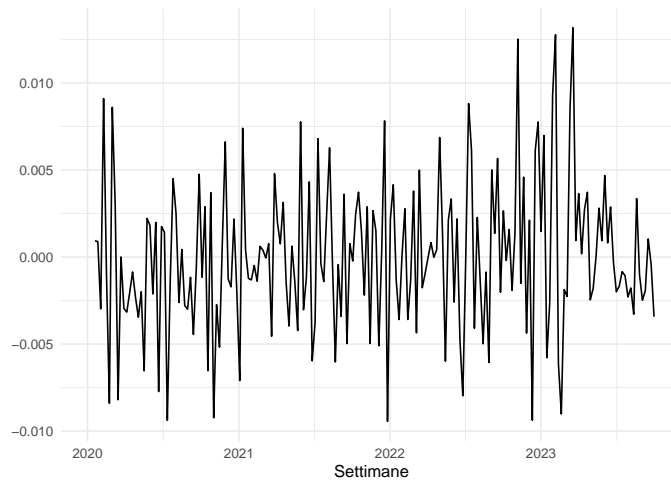


Figura 33: Grafico dei residui del modello LSTAR stimato, con variabile esterna Eurovita.

### 5.3.4 Stima del modello con *trend* deterministico

Il *trend* scelto è la curva *Gamma*, così come definita nel quarto capitolo. La stima è operata tramite la funzione di R “nls”, che implementa la Minimizzazione dei Quadrati non Lineari. Tale funzione permette di introdurre anche altri regressori nella stima, pertanto è possibile aggiungere informazione al modello per migliorare l’adattamento. Dato quanto appreso in precedenza sui diversi modelli, alcune possibili variabili indipendenti che si potrebbero inserire e testare sono, oltre alla componente autoregressiva di ordine 1, anche la serie stessa a ritardo 2 e 3, la serie dei BTP a ritardo 1 e 13, e in fine la serie Eurovita a ritardo 1.

Un problema però da affrontare è la scelta dell’istante di inizio del *trend*. Come accennato è possibile stimarla direttamente tramite il modello, ma non è possibile assicurare la correttezza della sua stima, né possono essere usati i test classici per la sua valutazione. Una ipotesi alternativa è quello di utilizzare le stime ricavate dai precedenti modelli sull’istante di transizione, in particolare le date stimate dai modelli TAR visti - valutati più corretti degli STAR - : il 2023 - 01 - 29 e il 2022 - 09 - 18.

Altro problema ancora è la selezione dei valori iniziali dei parametri. La funzione infatti richiede di fornire un vettore di condizioni iniziali da cui far partire l’algoritmo di ottimizzazione sottostante. La scelta è critica poiché determina fortemente i risultati e la convergenza dell’algoritmo, come già osservato. L’approccio preso è quello di fornire valori più o meno realistici, confrontandoli con quanto si può apprendere dall’andamento del tasso di riscatto e dai modelli precedenti.

Per cominciare si riporta il modello semplice così come presentato nel capitolo precedente. I valori iniziali scelti sono  $s = 2022-11-27$ , epoca intermedia tra le due proposte,



mentre gli altri sono  $\beta_0 = 0,02$ ,  $\beta_1 = 0,5$ , in funzione di quanto osservato dai modelli TAR,  $\beta_2 = 0.01$ ,  $\gamma_1 = 0.1$  e  $\gamma_2 = 1$ , ottenuti studiando diverse loro configurazioni per ottenere una curva *Gamma* compatibile con gli pseudo-dati. Nella tabella 14 sono raccolte le stime.

Tabella 14: Stima del modello con trend deterministico

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.01825	0.002698	6.765	1.608e-10
<b>b1</b>	0.6064	0.05788	10.48	1.467e-20
<b>b2</b>	9.23e-05	8.719e-05	1.059	0.2912
<b>g1</b>	0.08954	0.01802	4.969	1.497e-06
<b>g2</b>	2.441	0.4371	5.585	8.016e-08

Il parametro riferito al *trend* è stimato non significativo mentre i due parametri non lineari lo sono. Ciò è dovuto al ristretto range di valori assunto dalle osservazioni, che implica un valore a sua volta molto piccolo del parametro  $\beta_2$ . Infatti, andando a calcolare l'ordinata del punto di massimo della curva *Gamma*, ovvero  $\beta_2 \left(\frac{e\gamma_1}{\gamma_2}\right)^{-\gamma_2}$ , e supponendo un salto di circa 0,05 rispetto alla media del regime basso, mantenendo i valori stimati dei parametri non lineari, si può ricavare un valore realistico di  $\beta_2$  molto piccolo, vicino alla stima ottenuta, pari a<sup>117</sup>:

$$\beta_2 = \left(\frac{e\gamma_1}{\gamma_2}\right)^{\gamma_2} 0,05 = 1.7984464 \times 10^{-4} .$$

Valori molto piccoli o molto grandi sono difficili da stimare tramite questi algoritmi. Una soluzione possibile a questo è quello di riparametrizzare il modello, in modo tale che la dimensione del parametro viaggi in un *range* migliore. Ad esempio andando a fare una sostituzione di questo tipo:  $\beta_2 = \bar{\beta}_2^{\gamma_2}$ , si ottiene un modello con una stima significativa della variabile. Lo si può vedere osservando la tabella 15.

Tabella 15: Stima del modello con trend deterministico riparametrizzato

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.01825	0.002698	6.765	1.608e-10
<b>b1</b>	0.6064	0.05788	10.48	1.467e-20
<b>b2</b>	0.02224	0.006805	3.268	0.001286
<b>g1</b>	0.08955	0.01802	4.969	1.496e-06
<b>g2</b>	2.441	0.4371	5.585	8.014e-08

<sup>117</sup>Per il calcolo del punto di massimo si rimanda alle Appendici.

Tale trasformazione è però non necessaria e ciò lo si può vedere tramite la simulazione di una traiettoria del modello, riportata in figura 34.

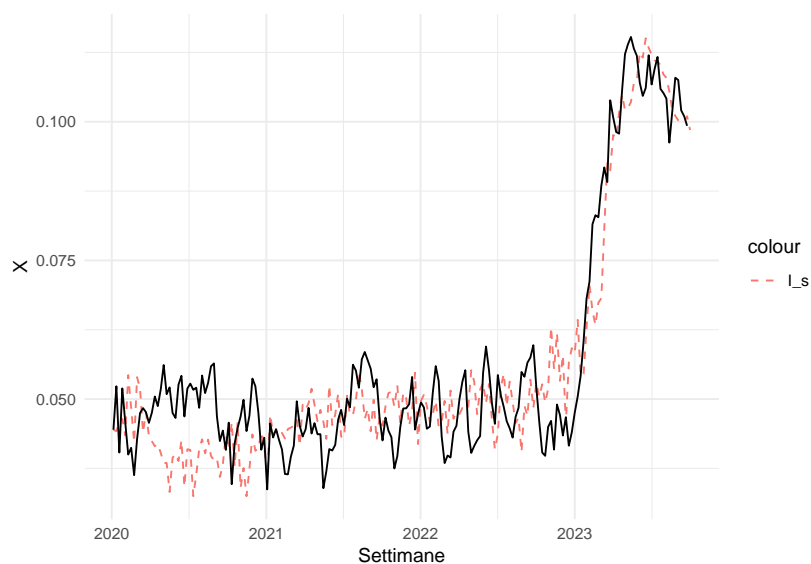


Figura 34: Modello a trend deterministico simulato.

La simulazione è buona, sebbene ci sia troppa volatilità, causata da una autocorrelazione residua. Si noti infatti come il modello utilizzato presenta un solo parametro autoregressivo, a dispetto dei precedenti. In figura 35 e 36 si possono vedere l'ACF e la PACF dei residui.

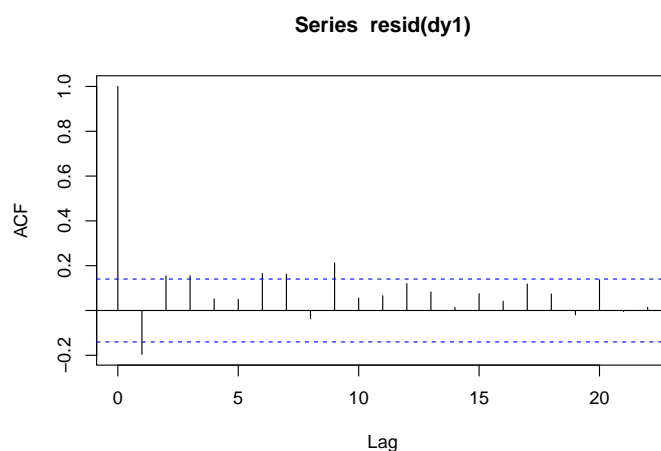


Figura 35: ACF dei residui del modello a trend deterministico.

Inserendo nel modello la variabile BTP - nella tabella 16 il parametro  $\mathbf{a1}$  - a un ritardo, questo risulta un regressore significativo, sebbene il suo parametro sia molto piccolo rispetto a  $\beta_1$ .

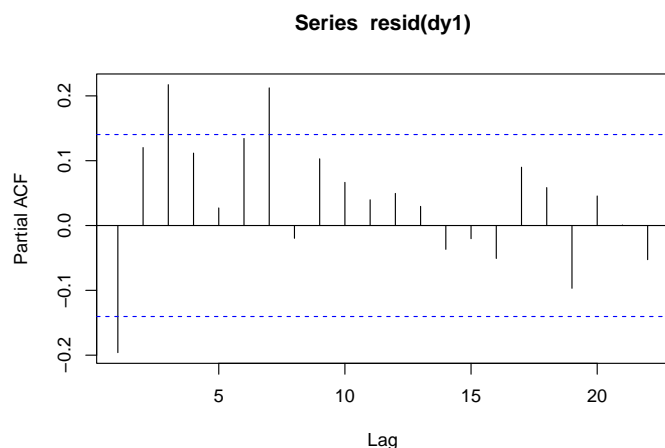


Figura 36: ACF dei residui del modello a trend deterministico.

Tabella 16: Stima del modello con trend deterministico e regressore BTP

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.02082	0.002743	7.588	1.437e-12
<b>b1</b>	0.5131	0.06286	8.162	4.542e-14
<b>b2</b>	1.957e-05	2.165e-05	0.9038	0.3672
<b>g1</b>	0.1144	0.0195	5.868	1.946e-08
<b>g2</b>	3.157	0.4977	6.344	1.61e-09
<b>a1</b>	0.1046	0.03041	3.439	0.0007182

Per quanto piccolo, risulta comunque più alto di quanto si otterrebbe stimando il modello senza il *trend* deterministico, e, ancora più rilevante, questo secondo modello risulta in generale peggiore del primo, a conferma del fatto che il picco dei riscatti non può essere colto da un semplice modello di regressione per le serie storiche con covariata i tassi di mercato. Le stime di questo secondo modello sono riportate in tabella 17.

Tabella 17: Stima del modello AR con regressore BTP

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.001479	0.0009762	1.516	0.1313
<b>b1</b>	0.9527	0.02199	43.32	4.655e-101
<b>a1</b>	0.06505	0.0312	2.085	0.03839

Inserendo invece la variabile riferita ad Eurovita, il suo parametro risulta non significativo; il motivo è legato a quanto osservato in precedenza su tale variabile.

Tornando al problema dell'autocorrelazione residua, per ridurla è possibile inserire anche il secondo ritardo della serie dei riscatti, così come era stato fatto nei modelli TAR precedenti. Ponendo  $\beta_2$  come parametro per il secondo ritardo e  $\beta_3$  per il *trend*, si ottengono le stime riportate nella tabella 18.

Tabella 18: Stima del modello con trend deterministico a due ritardi

	Estimate	Std. Error	t value	Pr(> t )
<b>b0</b>	0.01216	0.00285	4.267	3.137e-05
<b>b1</b>	0.4118	0.06902	5.967	1.182e-08
<b>b2</b>	0.327	0.06913	4.731	4.381e-06
<b>b3</b>	0.0001982	0.0001999	0.9917	0.3226
<b>g1</b>	0.08875	0.02246	3.952	0.0001098
<b>g2</b>	2.085	0.4874	4.278	3.003e-05

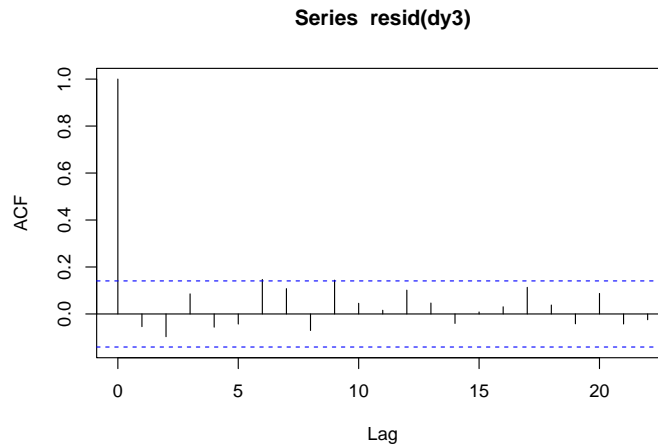


Figura 37: ACF dei residui del modello a trend deterministico.

Come anticipato, è possibile stimare anche la data di transizione dal modello stesso. La funzione però fatica a raggiungere il punto di ottimo: l'algoritmo converge ad un certo valore solo settando  $\gamma_2 = 1$  e il valore a cui converge dipende fortemente dalla scelta iniziale del parametro. Infatti, applicando la procedura di stima al modello con due ritardi, il valore dell'istante di transizione passa dal 2022-11-27 circa, partendo dalla data 2022-11-13, al 2023-02-12, partendo invece dal 2023-01-22. Quindi la procedura non è affidabile per la stima di tale data.

**Stima per la seconda e la terza HRG:** In fine, sono state operate le stime per le due serie rimanenti dello pseudo-campione. I modelli giudicati migliori ricalcano i

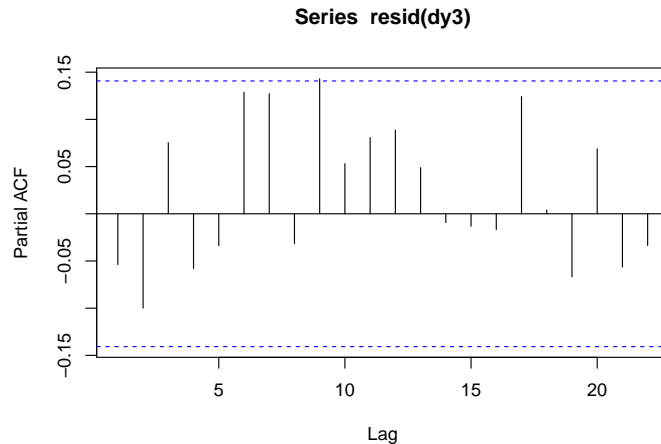


Figura 38: ACF dei residui del modello a trend deterministico.

risultati ottenuti per le stesse due serie nei modelli TAR. Infatti, alla seconda serie sono stati rimossi i primi 70 valori e il ritardo scelto per la serie stessa è stato pari a 1; così per la terza HRG si è utilizzato il ritardo 1.

Si vuole in fine ricordare ancora che questo modello a *trend* deterministico è stato costruito per poter rappresentare al meglio l'andamento teorico dell'evento di massa e per adattarsi al meglio al *trend* osservato. La sua utilità quindi risiede non nella sua capacità esplicativa del fenomeno, dato che non presenta la causa del picco, ma nella capacità di *fitting* della serie e nella sua capacità di adattamento e flessibilità nella sua costruzione.

### 5.3.5 La valutazione del *mass lapse risk*

Tramite i modelli ritenuti migliori, si vuole ora poter ottenere un possibile stima dell'impatto del *mass lapse* sui tre portafogli. La procedura è costruita per poter essere coerente con quanto indicato da *Solvency* in merito alla valutazione dei rischi, ovvero determinare la distribuzione del numero aleatorio riferito all'impatto e calcolarne il VaR. Nel caso della serie dei riscatti si procederà nel modo seguente: preso un modello, si andrà a simulare più volte il regime alto - associato all'evento estremo - , in modo tale da ottenere la distribuzione empirica del massimo della serie; tramite la FdR empirica si stimerà il VaR per il livello  $\alpha$  fornito dalla Direttiva, ovvero sia 99,5%.

Più precisamente, il massimo è riferito ad un certo periodo, è quindi necessario fissare un intervallo di tempo per il calcolo. Poiché la stessa Direttiva presuppone, nella

valutazione del rischio di riscatto di massa, che l'evento sia circoscritto ad un anno, si può utilizzare questo valore come lunghezza dell'intervallo da cui ricavare il massimo.

Prendendo come esempio il modello TAR con variabile esterna BTP, si riporta in figura 39 una traiettoria del regime alto, con il valore massimo segnato.

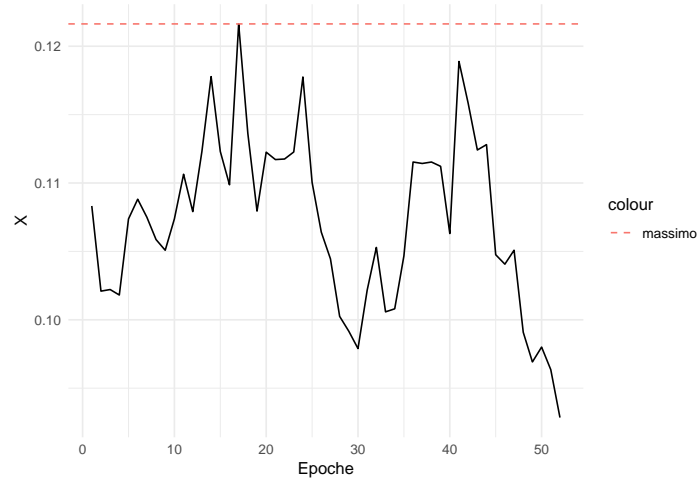


Figura 39: Simulazione regime alto del modello TAR stimato con variabile BTP.

Risimulando più volte, in questo caso il numero scelto è di 100.000 simulazioni, si ottiene la densità empirica, presentata in figura 40, del massimo su 52 settimane, da cui si può calcolare il VaR.

La procedura, così come descritta, non può essere applicata al modello con *trend* deterministico, poiché questo non presenta un “regime alto” come nel caso degli RSM, si deve perciò procedere in modo leggermente diverso. Quello che verrà fatto, per ottenere una valutazione coerente con le altre stime, sarà di simulare la traiettoria del processo supponendo che l'istante di transizione parta da inizio anno. La durata scelta della simulazione sarà tale da permettere al processo di raggiungere il suo picco, ricordando che la curva utilizzata per il *trend* ammette un solo massimo. Per i modelli stimati si è giudicata sufficiente una durata di 104 settimane. Tramite i valori simulati verrà poi calcolato il massimo.

Si applica la procedura a tutti i modelli ritenuti migliori tra quelli visti in precedenza, riassunti qui di seguito in punti:

1. TAR per BTP (**TY1** e **TZ1**):

- per la prima serie si ha ritardo  $d = 13$  settimane, ordine regime basso  $p_1 = 2$  e regime alto  $p_2 = 1$ ;
- per la seconda  $d = 6$ ,  $p_1 = 1$  e  $p_2 = 1$ , privo dei primi 70 valori;
- per la terza  $d = 1$ ,  $p_1 = 1$  e  $p_2 = 1$ ;

2. TAR per Eurovita (**TY2** e **TZ2**):

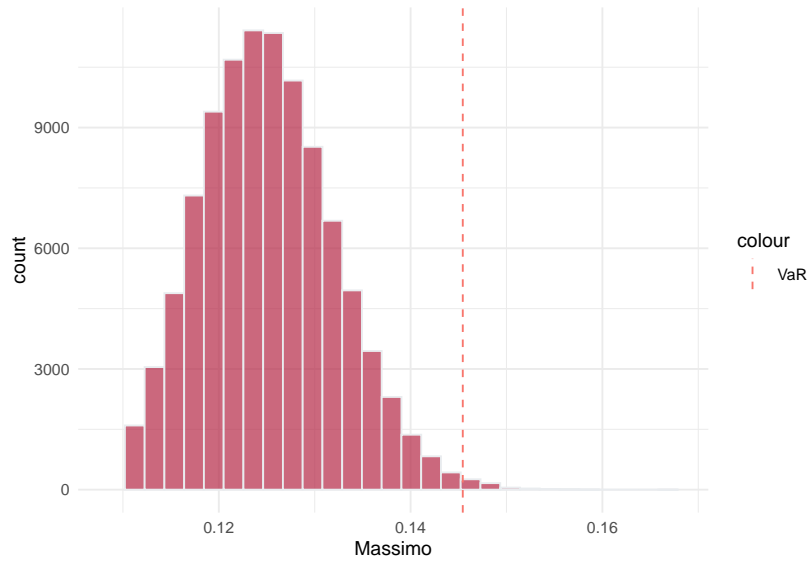


Figura 40: Densità empirica del massimo su 52 settimane del modello TAR stimato, con variabile BTP.

- solo la prima serie, con  $d = 1$  settimana,  $p_1 = 2$  e  $p_2 = 1$ ;

### 3. SETAR (**TY3** e **TZ3**):

- per la prima  $d = 3$ ,  $p_1 = 2$  e  $p_2 = 1$ ;
- per la seconda  $d = 2$ ,  $p_1 = 1$  e  $p_2 = 1$ , privo dei primi 70 valori;
- per la terza  $d = 1$ ,  $p_1 = 1$  e  $p_2 = 1$ ;

### 4. Modello a *trend* deterministico a due ritardi, nella tabella **DY3**.

- per la prima  $p_1 = 2$  e  $p_2 = 1$ ;
- per la seconda  $d = 2$ ,  $p_1 = 1$  e  $p_2 = 1$ , privo dei primi 70 valori;
- per la terza  $d = 1$ ,  $p_1 = 1$  e  $p_2 = 1$ ;

Fra parentesi è stato indicato il nome del modello utilizzato nella tabella, nello specifico **Y** indica la variabile non trasformata, mentre **Z** indica quella trasformata - al logit - .

In tabella 19 viene riportata la tabella dei VaR stimati:

Tabella 19: Stime dei VaR per i diversi modelli e per le tre HRG

	I_s	II_s	III_s
<b>TY1</b>	0.1454	0.09923	0.4862
<b>TZ1</b>	0.1648	0.1228	0.545
<b>TY2</b>	0.1403	NA	NA

	I_s	II_s	III_s
<b>TZ2</b>	0.1507	NA	NA
<b>TY3</b>	0.1268	0.1024	0.3833
<b>TZ3</b>	0.1303	0.1431	0.387
<b>DY3</b>	0.1287	0.09486	0.4024
<b>DZ3</b>	0.1555	0.089	0.455

Come si può vedere, per la prima HRG le stime sono approssimativamente univoche. In generale la prima serie presentava meno problematiche dovute al sua costruzione rispetto alle altre due e questo ha permesso di ottenere delle stime dei parametri maggiormente consistenti con i dati osservati. Prendendo invece la terza serie, questa presentava delle peggiori stime per la variabilità, che hanno influito sulla stima dei parametri del secondo regime nei TAR e ciò è possibile vederlo nei VaR stimati per tale serie. In particolare, per questa HRG, il regime alto dei modelli TAR con variabile esterna BTP presenta un parametro autoregressivo prossimo a uno e questo ha prodotto quelle stime molto elevate - tra 0,5 e 0,4, del VaR. Un altro elemento importante da notare è che le stime prodotte dal modello a *trend* deterministico sono in generale molto conservative, soprattutto con la seconda HRG. Questo può essere dovuto al fatto che la variabilità sul picco, per questi modelli, è data solamente dall'errore del modello, mentre negli altri casi la variabilità è data anche dalla componente autoregressiva, che in generale diventa più preponderante nel regime alto, dove il parametro per la parte autoregressiva è, per tutti i modelli RSM stimati, molto alto.

Un punto però che si vuole far notare è come le tre serie manifestino un massimo molto diverso, solamente la terza HRG presenta dei valori vicini, o persino superiori, al 40% proposto dal Regolamento, mentre le altre si trovano molto al di sotto. Ciò a supporto di quanto più volte espresso in questo presente scritto, ovvero che la stima fornita dalla formula standard non è in grado di rappresentare al meglio l'effettivo rischio di *mass lapse* per diversi portafogli.



## Conclusioni

In conclusione, la presente Tesi si è occupata di fornire un'analisi esaustiva del rischio di riscatto di massa, il *mass lapse risk*, così come definito all'interno della Direttiva europea *Solvency II*, a cui sono soggette le compagnie assicurative. In particolare, ha proposto un possibile approccio, tramite i *Regime Switching Models*, per la modellizzazione di tale rischio e ne ha verificato la bontà di adattamento su un campione osservato.

La ricerca è cominciata con una breve descrizione della Direttiva, mettendo a risalto gli elementi più rilevanti per l'argomento d'indagine. Si è passati poi a definire il diritto di riscatto e i diritti affini incorporati in alcune polizze del ramo Vita, a definire il metodo di valutazione del rischio secondo la Normativa e ad analizzare tutti le variabili da cui esso dipende. Con il terzo capitolo si è aperta la sezione fondamentale, ovvero quella riguardante la modellizzazione. È stata fatta una rassegna degli approcci più utilizzati per la modellizzazione della serie storica dei tassi di riscatto, il *lapse rate*, e sono stati poi presentati quattro articoli proponenti quattro metodologie distinte per la valutazione del *mass lapse risk*. Con il capitolo successivo è stato invece introdotto un ulteriore approccio, rispetto a quelli già presentati, mai proposto in letteratura, basato sull'utilizzo della famiglia dei modelli *Regime Switching*, in particolare tramite i TAR, i SETAR e i STAR. Oltre a questi è stato definito in aggiunta un semplice modello con *trend* deterministico, che riproducesse l'andamento previsto del *lapse rate* secondo la teoria. L'ultimo capitolo ha svolto una verifica empirica dei modelli presentati tramite un campione osservato di tre tassi di riscatto riferiti a tre HRG di una compagnia assicurativa. La stima dei modelli ha permesso di ottenere delle valutazioni del rischio di riscatto di massa, per un fissato livello di confidenza, da confrontare con il valore fornito dalla Direttiva.

I risultati hanno mostrato come i modelli TAR possano essere dei validi strumenti per la valutazione del rischio in esame, da poter quindi inserire nel modello interno di una compagnia assicurativa. Si è dimostrato poi come la serie dei tassi di mercato, qui rappresentata dai BTP a 10 anni, siano correlati con la serie dei riscatti tramite la relazione implicata dal modello TAR e possano pertanto essere utilizzati come predittori per un evento di estinzione anticipata di massa; mentre invece, il crollo di Eurovita, rappresentato dalla variabile *Google trend* della parola "Eurovita", si è dimostrato solo parzialmente responsabile per il picco osservato dei riscatti. Diversamente, i modelli STAR sono stati rigettati, a causa della loro scarsità di adattarsi bene ai dati. È stata costruita una metodologia per ottenere una stima dell'impatto del *mass lapse* confrontabile con quella fornita dalla Direttiva e si è potuto constatare che il valore 40% previsto dalla formula standard risulta inadeguato, sia in positivo che in negativo, per la determinazione del requisito di capitale richiesto per il sotto-modulo riferito al rischio indagato.

Tali risultati, per quanto positivi, presentano alcune criticità. Il campione originale

è stato modificato e completato utilizzando dati esterni, poiché ritenuto troppo ridotto per poter ottenere stime significative. La procedura di fusione delle due serie ha parzialmente compromesso la validità delle stime. Inoltre, la serie ottenuta presentava un unico picco, quando invece una valutazione più corretta necessitava di un campione su cui si siano osservati più picchi. Le funzioni utilizzate per operare la stima dei modelli non permettevano di introdurre più regressori, né gli RSM in sé permettono di inserire più processi esterni da cui far dipendere la transizione tra regimi: per la serie dei riscatti può essere ragionevole infatti supporre che più fenomeni concorrano a determinare il salto. In aggiunta, il modello TAR, così definito, non permette di stimare l'effetto causale tra la serie dei riscatti e il processo esterno. I modelli STAR sono stati rigettati, ma il motivo potrebbe essere dovuto alle modifiche operate sul campione e alla funzione utilizzata per stimarlo; richiede perciò una vagliatura più attenta. Non sono stati provati altri modelli tra gli RSM, come le forme vettorizzate, ad esempio i *Vector* TAR. In ultimo, l'approccio scelto valuta il rischio a livello di gruppo di polizze omogenee, ma in generale è difficile supporre non vi sia eterogeneità residua tra i contratti e ciò richiederebbe di modificare la procedura per poter inserire variabili riferite al contraente in sé, come l'età e il capitale assicurato.

# Appendici

In queste Appendici sono riportate alcune definizioni e dimostrazioni utilizzate nel presente Scritto. Gli argomenti trattati saranno:

- La Stazionarietà;
- I Modelli AR;
- Il massimo della curva Gamma del modello a *trend* deterministico.

## 5.4 La Stazionarietà

Si danno due definizioni di stazionarietà, una “forte” e una “debole”, a seconda di quanto sia restrittivo il requisito:

**Definizione (Stazionarietà Forte)** *Un processo stocastico  $\{Y_t; t \in \mathbb{Z}\}$  è fortemente stazionario quando soddisfa, per ogni  $n \in \mathbb{N}$ , per ogni  $(t_1, \dots, t_n) \in \mathbb{Z}^n$  e per ogni  $s \in \mathbb{Z}$ , la relazione:*

$$F_{Y_{t_1}, \dots, Y_{t_n}} = F_{Y_{t_1+s}, \dots, Y_{t_n+s}} ,$$

dove i due termini indicano la FdR congiunta di  $Y_{t_1}, \dots, Y_{t_n}$  e  $Y_{t_1+s}, \dots, Y_{t_n+s}$ .

**Definizione (Stazionarietà Debole)** *Un processo stocastico  $\{Y_t; t \in \mathbb{Z}\}$  è debolmente stazionario quando soddisfa, per ogni  $t \in \mathbb{Z}$ , le seguenti tre relazioni:*

- 1)  $E[Y_t] = \mu \in \mathbb{R}$  ,
- 2)  $0 < V[Y_t] = \sigma^2 < +\infty$  ,
- 3)  $COV[Y_t, Y_{t-k}] = \gamma_k \in \mathbb{R}$  , per ogni  $k = 0, 1, \dots$  ,

dove  $\gamma_k$  viene chiamata “Autocovarianza” ed è in funzione del solo ritardo  $k$ . Per  $k = 0$  si ha  $\gamma_0 = \sigma^2$

La prima definizione richiede che la valutazione probabilistica sia la medesima per ogni traslazione temporale, mentre la seconda richiede più semplicemente che i primi due momenti siano finiti e che l’autocovarianza dipenda solo dal ritardo considerato. Pertanto, la prima impone una condizione più restrittiva della seconda, la quale, a dispetto della prima, può essere verificata empiricamente. Si noti però che la prima non implica la seconda, poiché non sono richiesti momenti finiti, né la seconda implica

la prima. Un caso particolare riguarda i processi *Gaussiani*, per i quali le due definizioni coincidono.

Entrambe le definizioni hanno alla base l'idea che il processo sia stabile nel tempo, quindi i suoi momenti o la sua distribuzione non si modificano nel tempo. Ciò fa sì che studiare un qualsiasi segmento della serie porta a conoscere la struttura probabilistica anche delle parti del processo non osservate. In particolare, la stima temporale dei momenti, fatta su quel segmento osservato, risulta sensata. Poiché infatti di un processo stocastico non si osservano più realizzazioni, ma si dispone di una sola determinazione, si vuole poter utilizzare quella traiettoria come un campione da cui stimare i momenti e ciò è - in parte - possibile grazie alla stazionarietà; così come accade per un campione statistico di unità indipendenti, per il quale si ricerca l'ipotesi di identica distribuzione per avere le buone proprietà degli stimatori campionari<sup>118</sup>.

La perdita della stazionarietà può portare a diversi problemi, a seconda di come si manifesti la non stazionarietà stessa, ma più in generale si può dire che i test di ipotesi, gli intervalli di confidenza e le previsioni possono essere non affidabili, questo perché gli stimatori OLS possono avere distribuzioni non normali anche per campioni grandi. Più nello specifico, i test di significatività non sono corretti perché basati sull'ipotesi di distribuzione asintotica normale degli stimatori, i coefficienti possono risultare non corretti e si possono verificare fenomeni di correlazione spuria. Per questo è necessario studiare l'origine della non stazionarietà di una serie per poi o eliminarla tramite una trasformazione del processo o modificare il processo stesso per poterne tenere conto<sup>119</sup>.

I motivi che possono generare la non stazionarietà di una serie sono molteplici, si possono però citarne due: la presenza di un *trend*, deterministico o stocastico, o la presenza di un *break* strutturale. Vi sono certamente anche altri motivi, ma questi due in particolare sono presenti nella serie dei riscatti e l'intento dell'utilizzo dei modelli che verranno via via descritti ora è quello appunto di poter inserire nel modello base, che descrive il *lapse rate*, il *trend* e il *break* relativi al *mass lapse risk*, che determinano la non stazionarietà della serie<sup>120</sup>.

Due utili funzioni che è possibile definire per un processo stazionario in senso debole sono<sup>121</sup>: la “*Funzione di Autocorrelazione*” (ACF) e la “*Funzione di Autocorrelazione Parziale*” (PACF). Esse rappresentano due validi strumenti per l'analisi della autocorrelazione di una serie e in particolare sono utili per individuare il modello che più si adatta a tale serie. La prima è definita come l'autocorrelazione del processo in funzione del ritardo  $k = 0, 1, \dots$ :

$$\rho_k = CoR[Y_t, Y_{t-k}] = \frac{CoV[Y_t, Y_{t-k}]}{\sqrt{V[Y_t]V[Y_{t-k}]}} = \frac{\gamma_k}{\sigma * \sigma} = \frac{\gamma_k}{\gamma_0} ;$$

<sup>118</sup>Per avere la correttezza e la consistenza sono richieste anche altre proprietà, che qui non saranno discusse: l'“*Asintotica Incorrelazione*” e l'“*Ergodicità*”

<sup>119</sup>Sulle problematiche causate dalla non stazionarietà si faccia riferimento a Stock and Watson (2020).

<sup>120</sup>Sulle cause della non stazionarietà si faccia sempre riferimento a Stock and Watson (2020).

<sup>121</sup>Sono definibili anche per i processi stazionari in senso forte, ma è necessario aggiungere l'ipotesi di momento secondo finito.

la seconda invece considera l'autocorrelazione “parziale” in funzione del ritardo, ovvero l'autocorrelazione tra il processo e un suo ritardo al netto dell'effetto delle variabili riferite ai ritardi intermedi tra le due. Fissato un valore di  $k > 0$  il valore è di fatto pari al coefficiente  $\beta_k$  di un  $AR(k)$ , riferito al  $k$ -esimo ritardo.

## 5.5 I Modelli AR

**Definizione (Modello Autoregressivo)** Il Processo Autoregressivo di ordine  $p$  ( $AR(p)$ ), con  $p \in \mathbb{N}$ , è definito come un processo stocastico a tempo discreto  $\{Y_t, t \in \mathbb{Z}\}$ , con  $t$  parametro rappresentante il tempo, per il quale vale la relazione:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + \epsilon_t, \quad t \in \mathbb{Z}$$

dove il processo  $\{\epsilon_t, t \in \mathbb{Z}\}$  è il cosiddetto errore del modello, definito come un processo *White Noise* di media zero e varianza  $0 < \sigma_\epsilon^2 < +\infty$  e tale per cui vale:  $E[\epsilon_t | Y_{t-1}, Y_{t-2}, \dots] = 0$ ,  $\forall t \in \mathbb{Z}$ . Si ha che  $(\beta_0, \dots, \beta_p) \in \mathbb{R}^{p+1}$ , da cui si ha anche che  $Y_t$  ha supporto su tutto  $\mathbb{R}$ , per ogni  $t \in \mathbb{Z}$ <sup>122</sup>.

Con questo modello si intende rappresentare quelle dinamiche - in questo caso economiche - che dipendono dai suoi stessi valori passati, fino ad un certo ritardo. Si può notare che il processo prende la forma di una regressione lineare dove le covariate sono i ritardi della serie: è per questo motivo che è chiamato “auto-regressivo” e si può verificare inoltre che i parametri possono essere stimati facilmente tramite gli OLS. Se vale infatti l'ipotesi che  $Y_t$  sia incorrelata con i valori dopo il  $p$ -esimo ritardo e il processo è stazionario, si dimostra che gli stimatori OLS, per un campione sufficientemente grande, mantengono le loro proprietà, come esposto al paragrafo precedente<sup>123</sup>. È possibile ottenere delle stime anche tramite la massimizzazione della funzione di verosimiglianza, ma è richiesto di assegnare una distribuzione agli errori, tipicamente la *Normale*.

I modelli Autoregressivi possono non essere stazionari. Affinché si abbia la stazionarietà debole ad esempio di un  $AR(1)$  si deve verificare che sia  $|\beta_1| < 1$ , mentre per avere la stazionarietà in generale per un  $AR(p)$  è richiesta una condizione più complessa: ad ogni modello di ordine  $p$  è possibile associare un polinomio detto *polinomio caratteristico autoregressivo*, la stazionarietà è soddisfatta quando le radici di questo polinomio sono in modulo maggiore di 1. È possibile verificare tale condizione tramite un test statistico, come ad esempio il test aumentato *Dickey-Fuller*. Dato un  $AR(p)$

<sup>122</sup>Una “*White Noise*” è definita come un processo stocastico incorrelato, con media nulla e varianza finita non nulla

<sup>123</sup>La definizione di stazionarietà a cui qui si sta facendo riferimento è la seconda, quindi il particolare si sta supponendo che i momenti primo e secondo sono finiti.

stazionario è possibile determinare agevolmente il valore atteso  $E[Y_t] = \mu$ , costante per ogni  $t \in \mathbb{Z}$  proprio per la stazionarietà. Infatti:

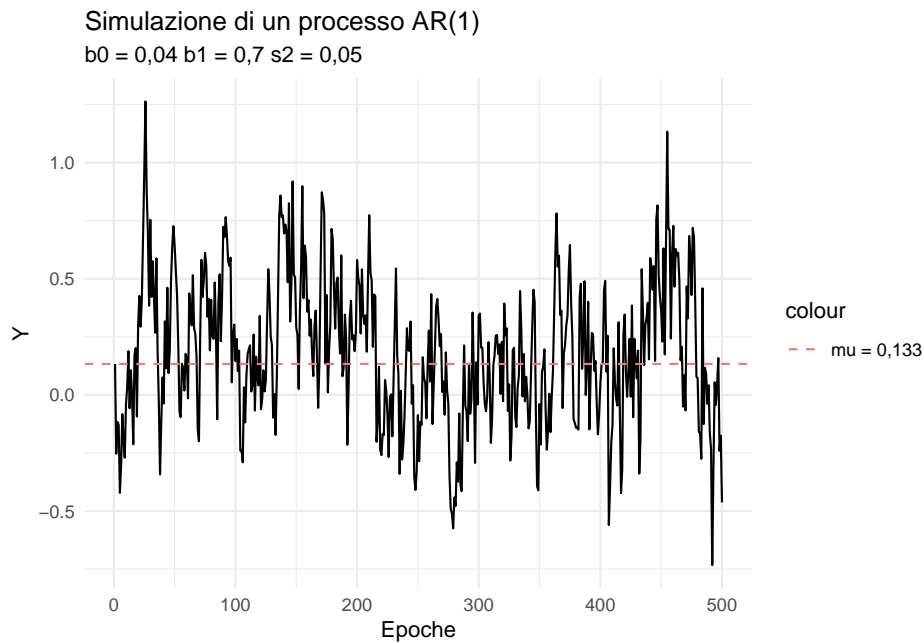
$$E[Y_t] = \beta_0 + \beta_1 E[Y_{t-1}] + \dots + \beta_p E[Y_{t-p}] + E[\epsilon_t] \text{ sse}$$

$$\mu = \beta_0 + \beta_1 \mu + \dots + \beta_p \mu + 0 \text{ sse}$$

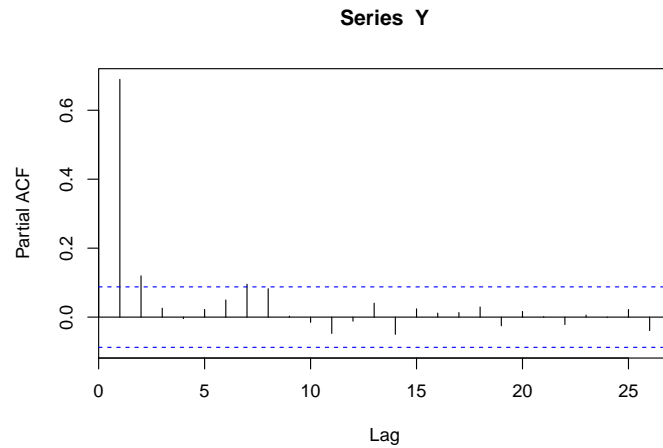
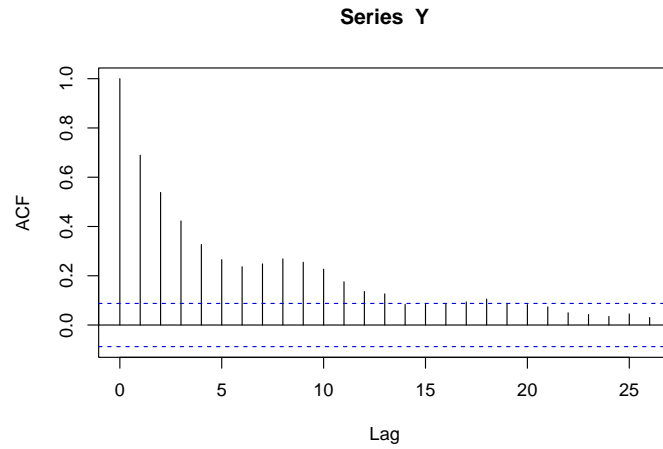
$$\mu = \frac{\beta_0}{1 - \beta_1 - \dots - \beta_p},$$

Da cui si può ricavare per l'AR(1) che  $\mu = \frac{\beta_0}{1 - \beta_1}$ .

Si riporta, per avere una informazione visiva sulla forma tipicamente assunta dai processi Autoregressivi, una simulazione di 500 valori da un AR(1) stazionario, con parametri:  $\beta_0 = 0,04$ ,  $\beta_1 = 0,7$  e  $\sigma_\epsilon^2 = 0,05$ . La linea tratteggiata indica la media del processo  $\mu$  calcolata sopra:



Un processo AR è caratterizzato inoltre da una particolare forma assunta dalla ACF e dalla PACF e ciò permette di identificare se un dato processo può essere un AR stazionario - condizione però solo necessaria - e quale ordine dovrebbe avere tale processo autoregressivo. In particolare, per un  $AR(p)$ , l'ACF è decrescente e convergente a zero, mentre la PACF ha  $p$  picchi come l'ordine del processo e i restanti valori nulli. Dato un campione della serie, è possibile stimare le due funzioni e rappresentarle tramite correlogramma, e tramite questi grafici valutare l'ipotesi che la serie sia un processo Autoregressivo con un certo ordine. Vengono riportati di seguito l'ACF e la PACF della serie simulata poco sopra:



I due correlogrammi presentano due rette blu indicanti un intervallo di confidenza. I valori compresi tra le barre sono considerabili non significati per un livello di confidenza del 95%. Dai grafici si può trarre quindi che la serie, come in effetti è, potrebbe essere un AR(1), secondo quanto detto sopra. Il correlogramma è molto utile anche per valutare una possibile non stazionarietà di una serie.

## 5.6 Il massimo della curva Gamma del modello a *trend* deterministico

La curva Gamma utilizzata nel modello con *trend* deterministico è stata definita in questo modo:

$$f(t) = \begin{cases} 0 & \text{per } t < S \\ \mu e^{-\beta(t-S)}(t-S)^{\alpha-1} & \text{per } t \geq S \end{cases}, \quad \mu, \beta, \alpha > 0.$$

Tale curva, si ricorda, è legata alla forma del comportamento teorico supposto della serie dei riscatti, ed è pertanto d'interesse lo studio del suo massimo.

Per  $\alpha < 1$  la funzione non presenta un massimo, ma ha un asintoto verticale in  $S$ , a più infinito. Quando invece  $\alpha = 1$  la funzione diventa la curva esponenziale, che ha massimo in  $S$ <sup>124</sup>. Caso più interessante è quello per  $\alpha > 1$ : la funzione risulta continua su tutto  $\mathbb{R}$  e derivabile ovunque tranne in  $S$ . Derivando la seconda componente della funzione, quindi per  $t > S$ , e ponendola uguale a zero si ha:

$$\begin{aligned} f'(t) &= \mu e^{-\beta(t-S)}(-\beta)(t-S)^{\alpha-1} + \mu e^{-\beta(t-S)}(\alpha-1)(t-S)^{\alpha-2} = \\ &= \mu e^{-\beta(t-S)}(t-S)^{\alpha-1} \left(-\beta + \frac{\alpha-1}{t-S}\right) = 0. \end{aligned}$$

Il primo fattore è positivo per ogni  $t > S$  - dati  $\mu, \beta, \alpha > 0$  - , il secondo invece si annulla per un unico valore:

$$t^* = \frac{\alpha-1}{\beta} + S,$$

e perciò risulta che  $f'(t^*) = 0$ . Si può verificare che per  $t < t^*$  la derivata è positiva mentre per  $t > t^*$  è negativa, quindi risulta che la curva è crescente prima di  $t^*$  e decrescente dopo  $t^*$  e pertanto  $t^*$  è punto di massimo.

Andando a calcolare la funzione  $f$  nel punto si ottiene:

$$\begin{aligned} y^* &= f(t^*) = \mu e^{-\beta\left(\frac{\alpha-1}{\beta} + S - S\right)} \left(\frac{\alpha-1}{\beta} + S - S\right)^{\alpha-1} = \\ &= \mu e^{-(\alpha-1)} \left(\frac{\alpha-1}{\beta}\right)^{\alpha-1} = \mu \left(\frac{e\beta}{\alpha-1}\right)^{-(\alpha-1)}. \end{aligned}$$

---

<sup>124</sup>In generale una funzione esponenziale non presenta un massimo, qui però ci si sta riferendo alla curva esponenziale definita nel quarto capitolo, la quale, nello specifico, è pari a zero prima di  $S$ .



## Ringraziamenti

Sono doverosi alcuni ringraziamenti. Per primo devo ringraziare il professor Milosovich che si è reso disponibile per farmi da relatore. Il suo aiuto nella comprensione degli argomenti teorici, le sue correzioni, i suoi consigli e la sua infinita pazienza sono stati fondamentali per la scrittura dell' Tesi. Provvidenziale è stato soprattutto il suo suggerimento di utilizzare i modelli *regime switching* per la modellizzazione del rischio di riscatto di massa.

Devo ringraziare infinitamente poi anche lo studio attuariale “*Visintin e Associati*” per avermi ospitato come tirocinante e per avermi fornito i dati essenziali per la valutazione empirica. In particolar modo devo ringraziare il dottor Roberto Bruni, attuario dell'ufficio, che mi ha fatto da tutor e correlatore della tesi. Le sue correzioni e le sue spiegazioni di argomenti tecnici della Direttiva e del funzionamento del modello interno di una compagnia assicurativa sono state indispensabili per la comprensione del fenomeno studiato, ma soprattutto è stato lui a propormi il tema della ricerca, il rischio di riscatto di massa.

Un ringraziamento speciale va al professor Carmeci per avermi dedicato tempo prezioso per approfondire alcuni aspetti teorici riguardanti le serie storiche non lineari e per avermi fornito importanti fonti bibliografiche sui modelli utilizzati.

Devo ringraziare ancora, di tutto cuore, i miei genitori, Roberto e Margherita, che hanno ecceduto nel loro ruolo di bravi genitori supportandomi oltre ogni ragionevolezza e investendo su di me con la stessa propensione al rischio di un giocatore d'azzardo. Spero di poterli rendere fieri di me come io sono fiero di loro.

Un ultimo grande ringraziamento è rivolto ad Alessia, non c'è persona che mi abbia dovuto sopportare più di lei. A te mi rivolgo direttamente: ti ringrazio di essermi stato vicino nei momenti disperati, sei tu la fonte da cui ho attinto le mie forze ed è a te che dedico tutte queste mie fatiche.

Grazie

## Bibliografia

- Azzone, Michele, Emilio Barucci, Giancarlo Giuffra Moncayo, and Daniele Marazzina. 2022. “A Machine Learning Model for Lapse Prediction in Life Insurance Contracts.” *Expert Systems with Applications* 191: 116261.
- Bacinello, Anna Rita, Enrico Biffis, and Pietro Millosovich. 2010. “Regression-Based Algorithms for Life Insurance Contracts with Surrender Guarantees.” *Quantitative Finance* 10 (9): 1077–90.
- Baione, Fabio, Davide Biancalana, and Paolo Angelis. 2021. “An application of Sigmoid and Double-Sigmoid functions for dynamic policyholder behaviour.” *Decisions in Economics and Finance* 44 (1): 5–22. <https://doi.org/10.1007/s10203-020-00279->.
- Barsotti, Flavia, Xavier Milhau, and Yahia Salhi. 2016. “Lapse Risk in Life Insurance: Correlation and Contagion Effects Among Policyholders’ Behaviors.” *Insurance: Mathematics and Economics* 71: 317–31. <https://doi.org/https://doi.org/10.1016/j.insmatheco.2016.09.008>.
- Biagini, Francesca, Tobias Huber, Johannes G. Jaspersen, and Andrea Mazzon. 2021. “Estimating Extreme Cancellation Rates in Life Insurance.” *Journal of Risk and Insurance* 88 (4): 971–1000. <https://doi.org/https://doi.org/10.1111/jori.12336>.
- Brémaud, Pierre. 2020. *Point Process Calculus in Time and Space*. Springer Cham. Doi. Vol. 10. Springer.
- Brockwell, Peter J, and Richard A Davis. 2002. *Introduction to Time Series and Forecasting*. Springer.
- CEIOPS’ Advice for Level 2 Implementing Measures on Solvency II: Technical Provisions Article 86 a Actuarial and Statistical Methodologies to Calculate the Best Estimate. 2009. CEIOPS Committee of european Insurance; Occupational Pensions Supervisors. <https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-L2-Final-Advice-on-TP-Best-Estimate.pdf>.
- CEIOPS-SEC-40-10 Solvency II Calibration Paper. 2010. CEIOPS Committee of european Insurance; Occupational Pensions Supervisors. <https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-Calibration-paper-Solvency-II.pdf>.
- Cerchiara, Rocco Roberto, A Gambini, RR Cerchiara, et al. 2009. “Multivariate Analyses to Modelling Lapse Risk Capital Charge Under Solvency II.” In *AFIR LIFE 2009*.
- Cheng, Chunli, Christian Hilpert, Aidin Miri Lavasani, and Mick Schaefer. 2023. “Surrender Contagion in Life Insurance.” *European Journal of Operational Research* 305 (3): 1465–79. <https://doi.org/https://doi.org/10.1016/j.ejor.2022.07.009>.
- Coles, Stuart, Joanna Bawa, Lesley Trenner, and Pat Dorazio. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Vol. 208. Springer.
- Considerations Regarding Dynamic Lapses in Actuarial Modeling. 2023. America Academy of Actuaries. <https://www.actuary.org/sites/default/files/2023-12/life->

- [paper-dynamic-lapses.pdf](#).
- Cox, Samuel H, and Yijia Lin. 2006. “Annuity Lapse Rate Modeling: Tobit or Not Tobit.” *Society of Actuaries*.
- Dar, Atul, and C Dodds. 1989. “Interest Rates, the Emergency Fund Hypothesis and Saving Through Endowment Policies: Some Empirical Evidence for the UK.” *Journal of Risk and Insurance*, 415–33.
- Dario Focarelli, Donatella Albano, Edoardo Marullo Reedtz. 2022. “TRENDS - Solvency - Anno III - Numero 07 - Luglio 2022.” ANIA Istituto per la Vigilanza sulle Assicurazioni. <https://www.ania.it/documents/35135/204382/ANIA+Trends+Solvency+-+Luglio+2022.pdf/8925a9a2-72e8-1247-a515-654fef74aed8?t=1659092769193>.
- Dickson, David CM, Mary R Hardy, and Howard R Waters. 2019. *Actuarial Mathematics for Life Contingent Risks*. Cambridge University Press.
- Dijk, Dick van, Timo Teräsvirta, and Philip Hans Franses. 2002. “Smooth Transition Autoregressive Models—a Survey of Recent Developments.” *Econometric Reviews* 21 (1): 1–47.
- DIRETTIVA 2009/138/CE DEL PARLAMENTO EUROPEO e DEL CONSIGLIO in Materia Di Accesso Ed Esercizio Delle Attività Di Assicurazione e Di Riasicurazione (Solvibilità II). 25 novembre 2009. Unione Europea. <https://eur-lex.europa.eu/eli/dir/2009/138/oj/ita>.
- Dobson, Annette J, and Adrian G Barnett. 2018. *An Introduction to Generalized Linear Models*. Chapman; Hall/CRC.
- Eling, Martin, and Dieter Kiesenbauer. 2014. “What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market.” *Journal of Risk and Insurance* 81 (2): 241–69.
- Fabio Baione, Andrea Santoro. 2019. “Sull’analisi Del Dynamic Policyholder Behaviour Nei Riscatti Attraverso Le Copule Bivariate.” [https://www.dss.uniroma1.it/it/system/files/pubblicazioni/Articolo\\_Baione\\_Santoro\\_06062022.pdf](https://www.dss.uniroma1.it/it/system/files/pubblicazioni/Articolo_Baione_Santoro_06062022.pdf).
- Franses, Philip Hans, and Dick Van Dijk. 2000. *Non-Linear Time Series Models in Empirical Finance*. Cambridge university press.
- Giroi, Federico, and Gary King. 2007. “Understanding the Lee-Carter Mortality Forecasting Method.” *Gking. Harvard. Edu*.
- Goel, Manish Kumar, Pardeep Khanna, and Jugal Kishore. 2010. “Understanding Survival Analysis: Kaplan-Meier Estimate.” *International Journal of Ayurveda Research* 1 (4): 274.
- Hull, John C, and Sankarshan Basu. 2016. *Options, Futures, and Other Derivatives*. Pearson Education India.
- Ingersoll, Jonathan E. 1987. *Theory of Financial Decision Making*. Vol. 3. Rowman & Littlefield.
- Kaas, Rob, Marc Goovaerts, Jan Dhaene, and Michel Denuit. 2008. *Modern Actuarial Risk Theory: Using r*. Vol. 128. Springer Science & Business Media.
- Kagraoka, Yusho. 2005. “Modeling Insurance Surrenders by the Negative Binomial Model.” In *JAFEE International Conference*. Vol. 1.

- Kiesenbauer, Dieter. 2012. “Main Determinants of Lapse in the German Life Insurance Industry.” *North American Actuarial Journal* 16 (1): 52–73.
- Kim, Changki. 2005. “Modeling Surrender and Lapse Rates with Economic Variables.” *North American Actuarial Journal* 9 (4): 56–70.
- Kim, Daegy. 2018. “Dynamic Lapse Modeling Using Korean Insurance Industry Data.” SOA Society of Actuaries. <https://www.soa.org/globalassets/assets/files/e-business/pd/events/2018/asia-pacific-symposium/asia-pacific-symposium-session-6b.pdf>.
- Koijen, Ralph S. J., Hae Kang Lee, and Stijn Van Nieuwerburgh. 2024. “Aggregate Lapsation Risk.” *Journal of Financial Economics* 155: 103819. <https://doi.org/https://doi.org/10.1016/j.jfineco.2024.103819>.
- Krolzig, Hans-Martin. 2002. “Regime-Switching Models.” <https://perhuaman.wordpress.com/wp-content/uploads/2014/09/krolzig2002.pdf>.
- Kuo, Weiyu, Chenghsien Tsai, and Wei-Kuang Chen. 2003. “An Empirical Study on the Lapse Rate: The Cointegration Approach.” *Journal of Risk & Insurance* 70 (September): 489–508. <https://doi.org/10.1111/1539-6975.t01-1-00061>.
- Liu, Xialu, Rong Chen, and Ruey Tsay. 2020. “NTS: An R Package for Nonlinear Time Series Analysis.” *The R Journal* 12 (2).
- Loisel, Stéphane, and Xavier Milhaud. 2011. “From Deterministic to Stochastic Surrender Risk Models: Impact of Correlation Crises on Economic Capital.” *European Journal of Operational Research* 214 (2): 348–57.
- Milhaud, Xavier, and Christophe Dutang. 2018. “Lapse tables for lapse risk management in insurance: a competing risk approach.” *European Actuarial Journal* 8 (1): 97–126. <https://doi.org/10.1007/s13385-018-0165-7>.
- Milhaud, Xavier, Stéphane Loisel, and Veronique Maume-Deschamps. 2011. “Surrender Triggers in Life Insurance: Classification and Risk Predictions.” *Bulletin Français d’Actuariat* 11 (January): 5–48.
- Modeling of Policyholder Behavior for Life Insurance and Annuity Products a Survey and Literature Review*. 2014. SOA Society of Actuaries. <https://www.soa.org/globalassets/assets/files/research/projects/research-2014-modeling-policy.pdf>.
- Outreville, J Francois. 1990. “Whole-Life Insurance Lapse Rates and the Emergency Fund Hypothesis.” *Insurance: Mathematics and Economics* 9 (4): 249–55.
- Paul Fell, Doug Scott e Robert Bugg. 2022. “Life Conference 2022 - Mass Lapse Reinsurance: Making It Stick.” Institute; Faculty of Actuaries. [https://www.actuaries.org.uk/system/files/field/document/C6\\_Robert%20Bugg\\_Paul%20Fell\\_Doug%20Scott.pdf](https://www.actuaries.org.uk/system/files/field/document/C6_Robert%20Bugg_Paul%20Fell_Doug%20Scott.pdf).
- REGOLAMENTO DELEGATO (UE) 2015/35 DELLA COMMISSIONE. 2014. Unione Europea. <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32015R0035>.
- Renshaw, Arthur E, and Steven Haberman. 1986. “Statistical Analysis of Life Assurance Lapses.” *Journal of the Institute of Actuaries* 113 (3): 459–97.
- Roberti, Roberto. 2017. *SOLVENCY II La Nuova Regolamentazione Prudenziiale SCR: Il Nuovo Requisito Patrimoniale Per Le Imprese Di As-*

- sicurazione*. IVASS Istituto per la Vigilanza sulle Assicurazioni. [https://www.ivass.it/pubblicazioni-e-statistiche/pubblicazioni/att-sem-conv/2017/sem-1907/Seminario\\_giornalisti\\_2017\\_Roberti.pdf](https://www.ivass.it/pubblicazioni-e-statistiche/pubblicazioni/att-sem-conv/2017/sem-1907/Seminario_giornalisti_2017_Roberti.pdf).
- Sax, Christoph, and Peter Steiner. 2013. “Temporal Disaggregation of Time Series.”
- Shaughnessy, Maureen, and Kevin Tewksbury. 2019. “U.s. Individual Life Insurance Persistency: A Joint Study Sponsored by the Society of Actuaries and LIM-RA.” SOA Society of Actuaries. <https://www.soa.org/globalassets/assets/files/resources/research-report/2019/2009-13-us-ind-life-persistency-update-report.pdf>.
- SOLVENCY II – LIFE INSURANCE*. 2016. Institute; Faculty of Actuaries. [https://www.actuaries.org.uk/system/files/field/document/IandF\\_SA2\\_SolvencyII\\_2016.pdf](https://www.actuaries.org.uk/system/files/field/document/IandF_SA2_SolvencyII_2016.pdf).
- SOLVENCY II La Nuova Regolamentazione Prudenziale Del Settore Assicurativo: Una Guida Semplificata*. 2016. IVASS Istituto per la Vigilanza sulle Assicurazioni. [https://www.ivass.it/pubblicazioni-e-statistiche/pubblicazioni/altre-pubblicazioni/2016/guida-solvency-ii/Guida\\_Solvency\\_II.pdf](https://www.ivass.it/pubblicazioni-e-statistiche/pubblicazioni/altre-pubblicazioni/2016/guida-solvency-ii/Guida_Solvency_II.pdf).
- Stefano Trionfetti, Dario Focarelli. 2024. “ANIA TREND Flussi e Riserve Al IV Trimestre 2023.” ANIA Associazione Nazionale fra le imprese Assicuratrici. [https://www.ania.it/documents/35135/671719/Newsletter+Vita+flussi+e+riskerve\\_IV+trim.2023.pdf/e2101cd2-2d45-03ca-a79a-4c1e46f8673c?version=1.1&t=1708602195111](https://www.ania.it/documents/35135/671719/Newsletter+Vita+flussi+e+riskerve_IV+trim.2023.pdf/e2101cd2-2d45-03ca-a79a-4c1e46f8673c?version=1.1&t=1708602195111).
- Stock, James H, and Mark W Watson. 2020. *Introduction to Econometrics*. Pearson.
- Tong, Howell. 1990. *Non-Linear Time Series: A Dynamical System Approach*. Oxford university press.
- Valvo, Paolo S. 2020. “A Bimodal Lognormal Distribution Model for the Prediction of COVID-19 Deaths.” *Applied Sciences* 10 (23). <https://doi.org/10.3390/app10238500>.
- Wüthrich, Mario V., and Michael Merz. 2008. *Stochastic Claims Reserving Methods in Insurance*.