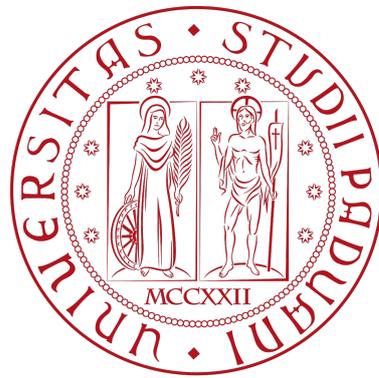


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE STATISTICHE

CORSO DI LAUREA MAGISTRALE IN  
SCIENZE STATISTICHE



**Allocazione di portafoglio a pesi dinamici e  
variabili predittive: un'applicazione al mercato  
Europeo**

*Relatore:*

PROF. MASSIMILIANO CAPORIN

*Laureando:*

CHIARA TONIOLO

2062597

Anno Accademico 2023/2024



# Indice

<b>Introduzione</b>	<b>4</b>
<b>1 Metodi e Criteri</b>	<b>9</b>
1.0.1 Funzione di utilità . . . . .	10
1.1 Metodo statistico parametrico a 2 passi . . . . .	10
1.1.1 Passo 1 . . . . .	11
1.1.2 Passo 2 . . . . .	13
1.2 Metodo non parametrico a 1 passo . . . . .	15
1.2.1 Funzione obiettivo . . . . .	16
1.2.2 Gradient boosting personalizzato . . . . .	17
1.3 Criteri di valutazione e riferimento . . . . .	21
1.3.1 Criteri di riferimento . . . . .	21
1.3.2 Criteri di valutazione . . . . .	22
<b>2 Dataset</b>	<b>25</b>
2.1 Indicatori aziendali ed economici . . . . .	27
2.2 Indicatori tecnici . . . . .	28
2.2.1 Indicatore Media Mobile . . . . .	29
2.2.2 Indicatore Momentum . . . . .	30
2.2.3 Indicatore sui volumi . . . . .	30
<b>3 Risultati</b>	<b>33</b>
3.1 Limite superiore $\omega^{max} = 1$ . . . . .	35
3.2 Limite superiore $\omega^{max} = 1.5$ . . . . .	42
3.3 Differenze prime dei predittori economici ed aziendali . . . . .	47
<b>Conclusione</b>	<b>49</b>

<b>A Alberi di regressione</b>	<b>55</b>
<b>B Scelta del numero di iterazioni M</b>	<b>59</b>
<b>C Analisi Indicatori aziendali ed economici</b>	<b>61</b>
<b>D Grafici</b>	<b>65</b>
<b>E Confronto grafici per <math>\omega^{max} = 1</math> e <math>\omega^{max} = 1.5</math></b>	<b>69</b>

# Introduzione

I metodi di previsione di allocazione ottima di portafoglio più diffusi nella finanza si basano comunemente su approcci a due fasi. In una prima fase utilizzano regressioni lineari per stimare i rendimenti azionari, i quali vengono successivamente impiegati per determinare i pesi ottimali per le decisioni di investimento.

Tuttavia, questi metodi sono stati oggetto di critica da parte di diversi studiosi, i quali sottolineano che l'approccio di ottimizzazione tramite minimi quadrati, pur generando previsioni statisticamente ottimali, non garantisce necessariamente un ritorno economico. Anche una lieve prevedibilità dei rendimenti può infatti avere un valore economico significativo, misurabile attraverso metriche basate sull'utilità e sul profitto. Questo ha portato a mettere in discussione l'affidabilità dei criteri puramente statistici, in quanto l'interesse primario degli investitori è il guadagno economico derivante dagli investimenti.

L'utilizzo di tali approcci a due fasi presuppone, inoltre, una conoscenza precisa del processo generatore dei rendimenti, che spesso risulta complesso e soggetto a possibili errori di specificazione. I modelli lineari, infine, sono restrittivi, poiché tendono a forzare i dati in una struttura predefinita, impedendo loro di seguire liberamente le proprie dinamiche.

Questo studio propone un metodo alternativo, empirico, non parametrico e flessibile, basato sulla massimizzazione diretta della funzione di utilità, il cui obiettivo è determinare i pesi di portafoglio in un unico passaggio, senza ricorrere alla stima dei rendimenti e senza imporre alcuna struttura ai dati.

Il vantaggio di ottenere i pesi direttamente in un unico passaggio consiste nell'evitare la stima della distribuzione condizionata dei rendimenti, un'operazione complessa e soggetta a errori. Questo approccio non parametrico consente di ridurre il rumore e minimizzare i potenziali errori di specificazione, creando un collegamento diretto tra le variabili predittive e i pesi di portafoglio.

glio, senza imporre strutture restrittive ai dati. Inoltre, permette di considerare relazioni non lineari tra le variabili, ampliando la flessibilità del modello.

La metodologia viene integrata con un algoritmo di apprendimento automatico, nello specifico un algoritmo di Gradient Boosting personalizzato, offrendo diversi vantaggi rispetto all'approccio parametrico per investitori con differenti preferenze di avversione al rischio, interessati a identificare quali variabili predittive e combinazioni di esse possano supportare al meglio le decisioni di allocazione.

Inoltre, questo nuovo metodo incorpora direttamente i limiti superiori e inferiori dei pesi di portafoglio tipicamente fissati a priori, integrandoli come parte della procedura. Al contrario, l'approccio a due fasi applica una troncatura solo a posteriori, dopo aver calcolato i pesi ottimali, per riportare i valori all'interno dell'intervallo stabilito. Tale operazione non fornisce spiegazioni sulle ragioni per cui alcune variabili risultano (o meno) predittori significativi per l'allocazione ottima. L'approccio diretto non parametrico, invece, aggira queste complicazioni, producendo risultati più facilmente interpretabili.

In economia e finanza, i metodi basati sull'apprendimento automatico sono stati spesso criticati per la mancanza di intuitività e interpretabilità, poiché generalmente orientati verso obiettivi specifici. La versione personalizzata che verrà introdotta, utilizza una funzione di utilità come funzione obiettivo, distinguendosi così dagli algoritmi statistici tradizionali e rendendo possibile un'interpretazione più chiara delle variabili predittive realmente significative per le decisioni di allocazione.

Questo approccio di discesa del gradiente personalizzato apprende iterativamente dagli errori di allocazione sul campione di stima, misurati tramite il gradiente della funzione obiettivo. In questo modo, fornisce previsioni ottimizzate dei pesi di portafoglio per il periodo successivo, riflettendo il comportamento tipico degli investitori, i quali aggiornano continuamente la loro strategia prima di prendere decisioni future.

Questo elaborato si focalizza sull'analisi del mercato europeo, impiegando variabili e indici specifici della zona Euro. Ci si concentra in particolare sul confronto tra le performance del metodo non parametrico con il classico approccio a due fasi, su un caso semplice di decisione tra un asset privo di rischio, rappresentato dal tasso *Euribor*, e un indice azionario rischioso, *l'Euro Stoxx 50*. Si

svolgono le previsioni dei pesi di portafoglio utilizzando un insieme di variabili predittive distinguibili in *Indicatori economici ed aziendali* e *Indicatori tecnici*. Si analizzano diversi profili di investitore per diverso grado di avversione al rischio, rappresentato dal parametro  $\gamma$ , e diversa quantità di capitale di investimento, rappresentato dal limite massimo dei pesi  $\omega^{max}$ . Vengono valutati i pesi stimati tramite criteri di valutazione economici quali, l'utilità media, il guadagno di rendimento certo equivalente (CER) e il rapporto Sharpe, i quali vengono confrontati con criteri di riferimento (benchmark) che rappresentano il potere predittivo derivante dalla "non informazione", costruiti partendo dalla stima dei pesi  $\omega_t$  ottenuta senza utilizzare i predittori.

Un confronto puntuale tra i due metodi mostra come quello diretto non parametrico porta a risultati migliori, più elevati rispetto al metodo basato sulla regressione lineare, in particolare utilizzando le variabili economiche ed aziendali. Le strategie di investimento valutate secondo i diversi criteri economici, utilizzando queste variabili, portano a risultati migliori e molto più elevati rispetto ai benchmark sia se analizzate in modo univariato che congiuntamente, per entrambi gli approcci. Gli indicatori tecnici, invece, mostrano generalmente una capacità predittiva limitata rispetto ai benchmark, con valori che spesso risultano inferiori a quelli di riferimento.

La struttura è organizzata come segue: nel Capitolo 1 vengono presentati nel dettaglio i due metodi di regressione lineare e di discesa del gradiente personalizzato e i criteri di riferimento e valutazione utilizzati per l'analisi e il confronto dei risultati finali, relativi a diverse tipologie di investitori, riportati nel Capitolo 3; infine, nel Capitolo 2 vengono descritte le variabili esplicative impiegate nella previsione dei pesi di portafoglio.



# Capitolo 1

## Metodi e Criteri

L'obiettivo principale di questa analisi è di ottenere il maggiore rendimento di portafoglio tramite un'allocazione ottimale che tiene conto del compromesso rischio-rendimento.

Viene considerato quindi un problema di decisione di allocazione per un investitore che ha come obiettivo la combinazione ottima di un portafoglio composto da un asset di mercato rischioso e uno privo di rischio, con un orizzonte temporale di un periodo. Dato un livello di ricchezza iniziale, l'ottimizzazione di tale problema si traduce nella ricerca dei pesi di portafoglio che massimizzano la funzione di utilità sottostante.

Sia  $\omega_t$  la variabile che simboleggia la porzione di valore di portafoglio allocata all'asset rischioso al tempo  $t$ , allora il rendimento del portafoglio al tempo  $t$  può essere rappresentato come segue:

$$\begin{aligned} r_{p,t} &= (1 - \omega_t)r_{f,t} + \omega_t r_{m,t} \\ &= r_{f,t} + \omega_t(r_{m,t} - r_{f,t}) \\ &= r_{f,t} + \omega_t r_{e,t} \end{aligned} \tag{1.1}$$

dove  $r_{e,t}$  indica l'eccesso di rendimento, dal periodo  $t - 1$  a  $t$ , ottenuto come differenza tra l'indice di mercato  $r_{m,t}$  e il tasso privo di rischio  $r_{f,t}$ .

L'ottimizzazione di (1.1) si otterrà quindi, mediante la determinazione della variabile  $\omega_t$  al tempo  $t$ , la quale verrà stimata utilizzando le informazioni fino al tempo  $t - 1$  contenute nei predittori  $\mathbf{x}_{t-1}$  ( $\omega_t \equiv \omega(\mathbf{x}_{t-1})$ ).

### 1.0.1 Funzione di utilità

Si considererà il caso di un investitore che vuole descrivere le proprie preferenze di investimento, tramite una funzione di utilità media-varianza di tipo quadratico, che tiene conto del compromesso tra profitto e rischio rappresentati rispettivamente dal valore atteso e dalla varianza del rendimento di portafoglio complessivo:

$$\begin{aligned}\max_{\omega_t} u(E_{t-1}(r_{p,t}), Var_{t-1}(r_{p,t})) &= \max_{\omega_t} \left\{ E_{t-1}(r_{p,t}) - \frac{1}{2}\gamma Var_{t-1}(r_{p,t}) \right\} \\ &= \max_{\omega_t} \left\{ E_{t-1}(r_{p,t}) - \frac{1}{2}\gamma\omega_t^2 Var_{t-1}(r_{e,t}) \right\} \quad (1.2)\end{aligned}$$

in cui  $\gamma > 0$  è chiamato *coefficiente di avversione al rischio* e rappresenta il grado del rischio, che l'investitore è disposto ad assumere e  $E_{t-1}(\cdot)$  e  $Var_{t-1}(\cdot)$  rappresentano rispettivamente, il valore atteso e la varianza condizionati ottenuti utilizzando le informazioni disponibili fino al tempo  $t - 1$ , contenute in un insieme di predittori. L'obiettivo sarà quindi la massimizzazione di (1.2) rispetto a  $\omega_t$  utilizzando il set informativo  $\mathbf{x}_{t-1}$ .

Tale funzione di utilità verrà utilizzata in entrambi i metodi analizzati in questo elaborato in modo tale da permettere un confronto dei risultati.

## 1.1 Metodo statistico parametrico a 2 passi

Il primo metodo analizzato è un metodo statistico definito "*a due passi*", basato sui primi due momenti dei rendimenti in eccesso  $r_{e,t}$ . Questo approccio consiste nella massimizzazione a priori della funzione di utilità (1.2), rispetto ai pesi  $\omega_t$ . Ponendo la derivata della funzione utilità uguale a zero, si ottiene la soluzione analitica dei pesi che massimizzano la funzione:

$$\begin{aligned}\omega_t^* &= \frac{E_{t-1}(r_{m,t}) - r_{f,t}}{\gamma Var_{t-1}(r_{e,t})} \\ &= \frac{E_{t-1}(r_{e,t})}{\gamma Var_{t-1}(r_{e,t})} \quad (1.3)\end{aligned}$$

Questa soluzione, è direttamente proporzionale al valore atteso del rendimento in eccesso al tasso privo di rischio; un incremento di tale valore suggerisce che l'investitore dovrebbe ottimizzare la propria strategia, aumentando l'esposi-

zione all'asset rischioso. Viceversa, è inversamente proporzionale rispetto alla varianza condizionata e al valore  $\gamma$ ; un aumento di questi valori, i quali rappresentano delle misure del rischio, suggeriscono invece di adottare un approccio più conservativo, riducendo la quota destinata all'investimento nell'indice di mercato.

### 1.1.1 Passo 1

Gli elementi ignoti dell'equazione (1.3) sono  $E_{t-1}(r_{e,t})$  e  $Var_{t-1}(r_{e,t})$ , i quali rappresentano rispettivamente, il valore atteso e la varianza condizionati della serie dei rendimenti in eccesso.

#### Valore atteso $E_{t-1}(r_{e,t})$

Il calcolo del primo elemento, che costituisce anche il numeratore dell'equazione, viene effettuato tramite una regressione lineare:

*Sia  $Z_{t-1} = (\mathbf{1}, \mathbf{x}_{i,1}, \dots, \mathbf{x}_{i,p})_{i=0}^{t-1}$  matrice dei predittori dove  $\mathbf{1}$  è un vettore unitario di intercetta ed  $\mathbf{x}_{i,j}$  con  $j = 1, \dots, p$  sono le serie storiche delle variabili esplicative osservate dal tempo  $i = 0$  fino a  $i = t - 1$ ; il modello di regressione lineare per i rendimenti in eccesso può venire rappresentato come segue:*

$$\mathbf{r}_{e,t} = Z_{t-1}\boldsymbol{\beta} + \varepsilon_t, \quad t = 1, \dots, T \quad (1.4)$$

*dove  $Z_{t-1}$  contiene una o più variabili esplicative e  $\boldsymbol{\beta} \in \mathbb{R}^p$  è il vettore dei coefficienti di regressione ignoti da stimare.*

I termini di disturbo  $\varepsilon_t$ ,  $t = 1, \dots, T$  si assumono indipendenti ed estratti da una distribuzione di media zero e varianza costante,  $\sigma_\varepsilon^2$ , e sono tali che  $E\{Z_{t-1}\varepsilon_t\} = 0$ .

I parametri ignoti  $\boldsymbol{\beta}$  vengono stimati mediante il criterio dei Minimi Quadrati, (*Ordinary Least Squares - OLS*). Questo approccio implica la ricerca del valore

di  $\boldsymbol{\beta}$ , che minimizza la somma dei quadrati dei residui

$$\hat{\boldsymbol{\beta}}_{OLS} = \min_{\boldsymbol{\beta}} \sum_{t=1}^T (r_{e,t} - Z_{t-1}\boldsymbol{\beta})^2 \quad (1.5)$$

La soluzione esplicita di tale problema di minimo è data da  $\hat{\boldsymbol{\beta}}_{OLS} = (Z_{t-1}^T Z_{t-1})^{-1} Z_{t-1}^T \mathbf{r}_{e,t}$ . Tale risultato viene utilizzato per ottenere le seguenti stime dei rendimenti in eccesso

$$\hat{r}_{e,t} = Z_{t-1} \hat{\boldsymbol{\beta}}_{OLS} \quad (1.6)$$

le quali potranno essere utilizzate come proxy per il numeratore dell'equazione dei pesi ottimi (1.3).

### **Varianza condizionata** $Var_{t-1}(r_{e,t})$

Per costruire il denominatore di (1.3) è necessario calcolare l'elemento ignoto  $Var_{t-1}(r_{e,t})$ .

In questa analisi viene utilizzato un approccio RiskMetrics in cui le varianze sono stimate dinamicamente mediante il modello a media mobile con decadimento esponenziale (*exponentially weighted moving average model - EWMA*) basato sui quadrati dei rendimenti in eccesso passati:

$$\begin{aligned} \sigma_t^2 &= (1 - \lambda) \sum_{s=1}^{t-1} \lambda^{s-1} r_{e,t-s}^2 \\ &= (1 - \lambda) r_{e,t-1}^2 + (1 - \lambda) \sum_{s=2}^{t-1} \lambda^{s-1} r_{e,t-s}^2 \\ &= (1 - \lambda) r_{e,t-1}^2 + \lambda \sigma_{t-1}^2 \end{aligned} \quad (1.7)$$

In questa espressione  $\sigma_t^2$  rappresenta la varianza al tempo  $t$ , calcolata sulla base delle informazioni disponibili al passo precedente  $t - 1$ , incorporate sia nella volatilità passata  $\sigma_{t-1}^2$  che nei quadrati dei rendimenti  $r_{e,t-1}^2$ .

I rendimenti vengono assunti con media nulla, mentre la costante  $\lambda$  rappresenta un fattore di liscio fissato, con valori compresi tra 0 e 1 ( $0 < \lambda < 1$ ). Per questa analisi viene utilizzato un valore  $\lambda = 0.95$ .

Questo parametro determina i pesi che vengono applicati ai rendimenti ed il modello è stato costruito in modo tale che, tutti i rendimenti storici contribuiscano alla stima della varianza, ma con un'influenza progressivamente

decescente, garantendo così che i più recenti abbiano un impatto maggiore.

La prima parte della rappresentazione ricorsiva di (1.7),  $(1 - \lambda)r_{e,t-1}^2$ , rappresenta *l'intensità della reazione* della volatilità agli eventi del mercato; al crescere di  $\lambda$  diminuisce la velocità di reazione all'informazione di mercato, racchiusa in  $r_{e,t-1}^2$ .

La seconda parte,  $\lambda\sigma_{t-1}^2$ , rappresenta invece la *persistenza* della volatilità. Sempre in questa rappresentazione, come valore iniziale della volatilità  $\sigma_0^2$  è possibile utilizzare la varianza campionaria, calcolata su una finestra di dati iniziali:  $\sigma_0^2 = Var(r_{e,1:k})$ .

Le dinamiche del modello esponenziale imitano da vicino quelle prodotte dal modello GARCH(1,1); più precisamente tale modello è simile nella forma al modello IGARCH,  $\sigma_t^2 = \omega + \alpha r_{e,t-1}^2 + \beta \sigma_{t-1}^2$ , dove  $\alpha + \beta = 1$ .

### 1.1.2 Passo 2

Avendo ora a disposizione tutti gli elementi necessari, è possibile ottenere i valori dei pesi di portafoglio empirici, effettuando il secondo passo del metodo, ovvero, sostituendo le quantità proxy per  $E_{t-1}(r_{e,t})$  e  $Var_{t-1}(r_{e,t})$ , calcolate al passo 1. Dato il prefissato coefficiente di avversione al rischio  $\gamma$ , l'equazione (1.3) viene riscritta come segue:

$$\hat{\omega}_t = \frac{\hat{r}_{e,t}}{\gamma \hat{\sigma}_t^2} = \frac{Z_{t-1} \hat{\beta}_{OLS}}{\gamma \hat{\sigma}_t^2} \quad (1.8)$$

L'espressione (1.8) riassume come i pesi di portafoglio possano essere determinati in funzione dei predittori  $\mathbf{x}_{t-1}$ , attraverso un processo in due fasi, basato sulla regressione predittiva lineare (1.4).

Questo primo metodo analizzato, sebbene di facile implementazione, presenta diverse complicazioni e criticità nelle performance finali del portafoglio:

1. Si considera il caso in cui la vendita allo scoperto non è ammessa e che quindi i pesi possono assumere solo valori non negativi  $\omega_t \geq 0$ ; questa assunzione, unita alla condizione principale sui pesi dove  $\sum_{i=1}^N \omega_i = 1$ , che nel nostro caso di soli due asset si traduce come  $(1 - \omega_t) + \omega_t = 1$ ,

può riassumersi come segue:

$$\omega_t \in [\omega^{min}, \omega^{max}] \quad (1.9)$$

dove il limite inferiore viene posto uguale a zero ( $\omega^{min} = 0$ ), impedendo, per l'appunto, la vendita allo scoperto. Per il limite superiore  $\omega^{max}$  possono venire considerate due diverse opzioni:  $\omega^{max} = 1.5$ , considerato da alcuni ricercatori vincolo realistico, il quale impedisce una leva finanziaria superiore al 50%. Un secondo valore comunemente utilizzato è dato da  $\omega^{max} = 1$ .

La procedura a due passi analizzata non garantisce che le stime dei pesi (1.8) ottenute, rispettino i limiti (1.9) e questo porta a dover effettuare un troncamento a posteriori dei risultati

$$\omega_t = \begin{cases} 0 & \text{se } \hat{\omega}_t \leq 0, \\ \hat{\omega}_t & \text{se } 0 < \hat{\omega}_t < \omega^{max}, \\ \omega^{max} & \text{se } \hat{\omega}_t \geq \omega^{max} \end{cases} \quad (1.10)$$

2. La troncatura delle stime a posteriori, necessaria per garantire il rispetto dei vincoli imposti, comporta una modifica sostanziale dei pesi stimati (1.8).

È importante sottolineare che tale operazione non fa parte del processo di regressione predittiva (1.4) e questo comporta che, tutte le interpretazioni statistiche legate al modello, come le statistiche test, le misure di bontà di adattamento e le conseguenti conclusioni e interpretazioni sui predittori  $\mathbf{x}_{t-1}$ , non sono più applicabili. Ad esempio, una variabile esplicativa, che risultava significativa nel modello, non necessariamente fornisce informazioni predittive utili ai fini delle decisioni di allocazione del portafoglio, dopo il troncamento. Analogamente, una variabile di scarsa significatività statistica potrebbe invece rivelarsi utile in tale contesto.

3. Una terza criticità risiede nella stima dei parametri  $\beta_i$ .

Si consideri il caso particolare della regressione lineare (1.4) con un solo predittore

$$r_{e,t} = \beta_{i,0} + x_{i,t-1}\beta_{i,1} + \varepsilon_{i,t} \quad (1.11)$$

In questo contesto, è di particolare interesse esaminare la prevedibilità dei rendimenti tramite il test di ipotesi nulla  $\beta_{i,1} = 0$ , basato sulla statistica  $t$ . Tuttavia, è noto che questo test di significatività può soffrire alla cosiddetta *distorsione di Stambaugh*, la quale si manifesta quando il predittore è altamente persistente e correlato con le innovazioni dei rendimenti  $\varepsilon_{i,t}$ , come riportato da Stambaugh (1999), Rapach e Zhou (2013), e Demetrescu et al. (2020).

4. Infine, l'approccio a due passi basato sulle regressioni predittive appartiene ai metodi plug-in. Sebbene il criterio dei minimi quadrati conduca a una soluzione ottimale dal punto di vista statistico, questo non sempre riflette le reali preferenze di un investitore, nell'allocazione ottimale di portafoglio. Pur essendo un metodo semplice per ottenere le stime necessarie al calcolo di (1.8), di solito non garantisce il raggiungimento dell'obiettivo finale, ossia la massimizzazione della funzione di utilità.

## 1.2 Metodo non parametrico a 1 passo

Per superare le limitazioni del metodo parametrico, legate alla regressione lineare predittiva, viene introdotto un secondo approccio empirico più flessibile, non parametrico e non lineare. L'obiettivo di questo metodo rimane la stima dei pesi di portafoglio ottimi  $\omega_t$  e, come nel caso precedente, si basa sulla massimizzazione della funzione di utilità.

L'ottimizzazione viene eseguita utilizzando un algoritmo di boosting, modificato specificatamente per adattarsi al nostro caso di studio.

La principale differenza rispetto al metodo precedente è data dal fatto che, in questo caso, non ci si concentra sulla distribuzione dei rendimenti in eccesso. I pesi  $\omega_t$ , invece, vengono modellati con una funzione diretta, potenzialmente non lineare, delle variabili esplicative  $\mathbf{x}_{t-1}$ , massimizzando la funzione di utilità dell'investitore.

Questo collegamento diretto incorpora tutte le informazioni rilevanti fornite dai predittori, indipendentemente dal fatto che tali variabili influenzino o meno la prevedibilità statistica dei rendimenti, evitando così le complicazioni 2 e 3 della regressione lineare nel metodo a due passi.

### 1.2.1 Funzione obiettivo

Questo secondo approccio verrà applicato utilizzando lo stesso set di dati di stima  $\{x_{t-1}, r_{e,t}\}_{t=1}^T$  impiegato nel precedente metodo dei minimi quadrati, dove  $r_{e,t}$  incorpora implicitamente le informazioni necessarie sia su  $r_{m,t}$  che su  $r_{f,t}$ .

Come primo passo il problema di massimizzazione della funzione di utilità (1.2) viene riformulato in un problema di minimizzazione

$$\min_{\omega_t} \left\{ \frac{1}{T} \sum_{t=1}^T -u_t \right\} = \min_{\omega_t} \left\{ \frac{1}{T} \sum_{t=1}^T - \left( r_{p,t} - \frac{1}{2} \gamma \omega_t^2 \sigma_t^2 \right) \right\} \quad (1.12)$$

dove  $r_{p,t}$  rappresentano i rendimenti di portafoglio risultanti (1.1) e  $\sigma_t^2$  la varianza condizionata, calcolata come nel metodo precedente in modo tale da consentire un confronto diretto dei risultati.

Anche in questo caso, vengono escluse le vendite allo scoperto, impostando quindi il limite inferiore in (1.9) a  $\omega^{min} = 0$ . Per rispettare il limite superiore  $\omega^{max}$  e superare la criticità 1 relativa al troncamento a posteriori del metodo a due passi, questo limite viene introdotto in modo esplicito nella procedura di stima dei pesi come segue:

$$\omega_t = \omega^{max} \lambda_t \quad (1.13)$$

dove

$$\lambda_t = \frac{1}{1 + \exp\left(-\frac{1}{\gamma \sigma_t^2} F(\mathbf{x}_{t-1})\right)} \quad (1.14)$$

Pertanto, il peso  $\omega_t$  viene principalmente determinato dalla funzione logistica  $\lambda_t$ , che, in combinazione con (1.13), garantisce che i pesi stimati rimangano sempre entro i confini definiti, ovvero  $\omega_t \in [0, \omega^{max}]$ .

L'equazione dei pesi non si basa più sulla stima dei rendimenti, ma sfrutta direttamente le informazioni contenute nei predittori  $\mathbf{x}_{t-1}$  attraverso la componente  $F(\mathbf{x}_{t-1})$ , che rappresenta una funzione complessa e non parametrica delle variabili esplicative. L'effetto di questa funzione in (1.14) viene regolato, come nel metodo di regressione precedente, dalle misure di rischio  $\sigma_t^2$  e  $\gamma$ .

È possibile riscrivere (1.14) utilizzando la seguente trasformata logistica:

$$\log\left(\frac{\lambda_t}{1-\lambda_t}\right) = \log\left(\frac{\omega^{max}\lambda_t}{\omega^{max}(1-\lambda_t)}\right) = \frac{F(\mathbf{x}_{t-1})}{\gamma\sigma_t^2} \quad (1.15)$$

Questa espressione consente un confronto diretto con la formulazione dei pesi in (1.8).

Tale formulazione rappresenta il logaritmo del rapporto delle probabilità (log odds) rispetto a  $\lambda_t$ . Valori elevati di  $\lambda_t$  corrispondono a un'alta probabilità di un evento favorevole, suggerendo così un incremento del valore del peso  $\omega_t$ .

## 1.2.2 Gradient boosting personalizzato

L'elemento centrale su cui si focalizza questo approccio è la determinazione della componente  $F(\mathbf{x}_{t-1})$  in (1.14). È su tale componente che il metodo applica l'algoritmo di gradient boosting personalizzato, tenendo conto delle complicazioni 1-4 del metodo *a 2 passi*.

Il boosting è una tecnica di apprendimento automatico che si basa sull'idea di utilizzare dei modelli di base semplici e deboli, detti *base learner*, e di combinare i loro risultati tramite una procedura iterativa, per creare un modello complessivo più robusto, in grado di fornire previsioni più accurate rispetto a quelle ottenute da un singolo modello.

Questo algoritmo ha ottenuto un grande successo in vari settori, tra cui la robotica, la statistica medica e l'economia. Anche in ambito finanziario esistono numerosi esempi di applicazione, come la previsione dei rendimenti azionari (Rossi e Timmermann (2015), Rossi (2018)), la previsione della volatilità (Mittnik, Robinzonov e Spindler (2015)), la modellazione della curva dei rendimenti (Audrino e Trojani (2007)) e i fallimenti nel settore bancario (Carmona, Climent e Momparler (2019)).

Il gradient boosting presenta diversi vantaggi. In primo luogo, esegue un processo, parallelo a quello di stima, che può essere interpretato come una selezione del modello: ad ogni iterazione, viene costruito un nuovo modello (funzione di base) con l'obiettivo di correggere gli errori commessi dal modello precedente, ottimizzando una funzione di perdita selezionata. Questo permet-

te al modello di adattarsi in modo dinamico ai dati.

L'uso di modelli di base non parametrici consente inoltre di catturare relazioni non lineari tra i predittori.

Un ulteriore vantaggio può derivare dalla scelta appropriata del modello di base. Se questo modello include un meccanismo per selezionare i predittori più rilevanti, l'algoritmo di boosting realizza una selezione implicita delle variabili a ogni iterazione. Ciò permette di identificare quali variabili predittive apportano un contributo più significativo, minimizzando, così, il rischio di sovradattamento ai dati e migliorando la robustezza del modello rispetto ai dati futuri.

In altre parole, l'algoritmo sarebbe in grado di selezionare di volta in volta le variabili rilevanti, scartando quelle non significative. Questo lo rende particolarmente adatto per dataset con molti predittori, anche quando il loro numero supera quello delle osservazioni, e specialmente in presenza di predittori correlati tra loro, risolvendo il problema della multicollinearità, dove i metodi lineari tendono a fallire.

Il contesto in cui si svolge la nostra analisi richiede un'impostazione diversa rispetto a quella tradizionale dell'algoritmo di gradient boosting. Normalmente, questi algoritmi si basano su approcci puramente statistici, con procedure di aggiornamento che utilizzano l'errore quadratico medio (MSE) o la verosimiglianza. Nel nostro caso, però, la funzione di perdita viene sostituita da una diversa funzione obiettivo: la funzione di utilità negativa, come descritto in (1.12).

Questa modifica evidenzia come un metodo meccanico di apprendimento automatico, come il gradient boosting, possa essere personalizzato e reso più flessibile e allineato al nostro obiettivo di ottimizzazione delle decisioni di allocazione di portafoglio.

Seguendo Bühlmann e Hothorn (2007) e la maggior parte degli stimatori boosting, il nostro algoritmo può essere descritto come segue:

**Input:** I dati  $\{\mathbf{x}_{t-1}, r_{e,t}\}_{t=1}^T$   
 Il coefficiente di avversione al rischio  $\gamma$   
 La proxy di volatilità selezionata  $\sigma_t^2$   
 Il numero di iterazioni  $M$

1. Inizializzazione dell'algoritmo con un valore costante:  $F_0(\mathbf{x}_{t-1}) = 0$

2. per  $m = 1, \dots, M$ :

a) calcolo degli pseudo-residui:

$$\tilde{y}_{t,m} = - \left[ \frac{\partial(-u_t)}{\partial F(\mathbf{x}_{t-1})} \right]_{F(\mathbf{x}_{t-1})=F_{m-1}(\mathbf{x}_{t-1})} \quad \text{per } t = 1, \dots, T \quad (1.16)$$

il gradiente si ottiene tramite la regola della catena:

utilizzando l'espressione (1.14), data la proxy della volatilità selezionata  $\sigma_t^2$  e notando il segno negativo in (1.12), otteniamo:

$$\begin{aligned} \frac{\partial(-u_t)}{\partial F(\mathbf{x}_{t-1})} &= \frac{\partial(-u_t)}{\partial \omega_t} \frac{\partial \omega_t}{\partial \lambda_t} \frac{\partial \lambda_t}{\partial F(\mathbf{x}_{t-1})} \\ &= -(r_{e,t} - \gamma \omega_t \sigma_t^2) \frac{\omega^{max} \lambda_t (1 - \lambda_t)}{\gamma \sigma_t^2} \end{aligned} \quad (1.17)$$

b) adattamento del modello di base (base learner)  $h_m(\mathbf{x}_{t-1})$  ai dati  $\{\mathbf{x}_{t-1}, \tilde{y}_{t,m}\}_{t=1}^T$ . Si utilizza un modello di regressione semplice con bassa varianza, come un albero stump o al più a due livelli.

c) aggiornamento delle previsioni:

$$F_m(\mathbf{x}_{t-1}) = F_{m-1}(\mathbf{x}_{t-1}) + \nu h_m(\mathbf{x}_{t-1})$$

**3. Output:**  $F_M(\mathbf{x}_{t-1})$

l'output  $F_M(\mathbf{x}_{t-1})$  fornisce la soluzione al problema di minimizzazione descritto in (1.12) utilizzando ad ogni iterazione le equazioni dei pesi (1.13) e di  $\lambda$  (1.14).

Vengono utilizzati gli stessi dati  $\{\mathbf{x}_{t-1}, r_{e,t}\}_{t=1}^T$  su cui viene adattato il modello

di regressione nell'approccio *a 2 passi*, al fine di consentire un confronto diretto tra i risultati ottenuti.

L'algoritmo appena descritto è stato definito da Bühlman e Hothorn (2007) e Friedman (2001) come *Algoritmo di discesa del gradiente funzionale* (*Functional Gradient Descent - FGD*). Questa specificazione rappresenta una variante dei classici algoritmi di boosting e può essere interpretata come una tecnica di ottimizzazione nello spazio delle funzioni tramite discesa del gradiente.

Dopo aver inizializzato la procedura con un valore scalare, a ogni iterazione vengono calcolati gli pseudo-residui come il gradiente negativo della funzione di utilità  $-u_t$ , valutato nella funzione  $F_{m-1}(\mathbf{x}_{t-1})$  ottenuta al passo precedente  $m - 1$ . Successivamente, un modello di base  $h_m(\mathbf{x}_{t-1})$  viene adattato ai dati, utilizzando come variabile risposta le quantità calcolate al punto 2(a).

L'aggiornamento della stima  $F_m(\mathbf{x}_{t-1})$  avviene attraverso il modello di base, il cui contributo viene moderato da un tasso di apprendimento  $\nu \in [0,1]$ , che regola la velocità di aggiornamento in modo tale da prevenire il sovradatamento. In linea con studi precedenti sul boosting, il valore di  $\nu$  viene fissato a 0.1.

L'intero processo descritto nel punto 2 viene ripetuto per  $M$  iterazioni, portando gradualmente al miglioramento della stima finale  $F_M(\mathbf{x}_{t-1})$

Gli aspetti chiave che assicurano le migliori prestazioni dell'algoritmo sono il numero di iterazioni  $M$  e il modello di base utilizzato nel passo 2. I dettagli relativi al modello saranno presentati e approfonditi nell'Appendice A. Per quanto riguarda il numero di iterazioni  $M$ , questo valore viene fissato a 500 ( $M = 500$ ); verrà mostrato nell'Appendice B che un ulteriore aumento delle iterazioni non apporta un significativo miglioramento dei risultati rispetto all'aumento del costo computazionale impiegato.

L'output generato dall'algoritmo di gradient boosting personalizzato  $F_M(\mathbf{x}_{t-1})$  applicato a  $\lambda_t$  in (1.14) e successivamente a  $\omega_t$  in (1.13) consente di stimare i pesi di portafoglio ottimali.

Questa stima viene ottenuta massimizzando direttamente la funzione di utilità  $u_t$  e affrontando le criticità indicate nei punti 1-4, che affliggevano il metodo

basato sulla regressione lineare.

## 1.3 Criteri di valutazione e riferimento

Per valutare l'efficacia dei nostri metodi di determinazione dei pesi di portafoglio, è importante introdurre i criteri principali di valutazione e gli approcci di riferimento utilizzati. Mentre la ricerca passata ha spesso dato priorità ai criteri statistici e alla prevedibilità dei rendimenti, si ritiene più rilevante misurare il valore economico delle diverse strategie. Pertanto, l'attenzione non verrà concentrata sulla significatività statistica dei predittori, ma sulla loro reale utilità nel migliorare le decisioni di asset allocation.

### 1.3.1 Criteri di riferimento

Questi parametri vengono costruiti basandosi sull'assunzione che le variabili esplicative  $\mathbf{x}_{t-1}$  non posseggano alcun potere predittivo nella determinazione dei pesi di portafoglio. Per i due metodi i pesi vengono costruiti come segue

#### Approccio a 2 passi

Nella regressione predittiva lineare si utilizza un valore costante nel tempo del rendimento in eccesso  $\bar{r}_{e,t} = \frac{1}{t} \sum_{j=1}^t r_{e,j}$ , chiamato *Media storica (Historical Average - HA)*.

Di conseguenza, nell'approccio a due passi l'equazione dei pesi (1.8) si riscrive come

$$\bar{\omega}_t = \frac{\bar{r}_{e,t}}{\gamma \hat{\sigma}_t^2} \quad (1.18)$$

si noti come nonostante il rendimento atteso sia costante, i pesi  $\bar{\omega}_t$  variano nel tempo grazie alla stima della varianza condizionata  $\hat{\sigma}_t^2$ .

#### Approccio a 1 passo

Nell'approccio a 1 passo i pesi  $\hat{\omega}_t$  vengono stimati utilizzando come variabile predittiva solo il valore costante  $\mathbf{x}_{t-1} = \mathbf{1}$ . La stima della componente  $F(\mathbf{x}_{t-1})$ , nell'algoritmo di gradient boosting personalizzato, sarà quindi guidata solamente dalla volatilità condizionata  $\hat{\sigma}_t^2$ .

I pesi stimati in questo modo vengono valutati tramite i criteri di valutazione che seguiranno a questo paragrafo e i valori risultanti saranno utili a valutare se un variabile possiede o meno potere predittivo utile.

Intuitivamente, se la variabile  $\mathbf{x}_{t-1}$  contenesse informazioni valide alla previsione dei pesi allora  $\hat{\omega}_t$  dovrebbe fornire risultati migliori rispetto a  $\bar{\omega}_t$ .

### 1.3.2 Criteri di valutazione

#### Utilità media

Il primo criterio per valutare le performance dei nostri risultati è l'utilità media, una misura che riflette il compromesso tra rischio e rendimento. Questo indicatore si basa sulla funzione di utilità dell'investitore, ottimizzata nei due approcci (1.2) e (1.12). Il calcolo dell'utilità media si ottiene sommando i rendimenti del portafoglio, opportunamente penalizzati per il rischio, fornendo così una valutazione che tiene conto sia delle opportunità di profitto sia della loro esposizione al rischio

$$\bar{u}(\hat{\omega}) = 100 \times \frac{1}{T} \sum_{t=1}^T \hat{u}_t = \frac{100}{T} \sum_{t=1}^T \left\{ \hat{r}_{p,t} - \frac{1}{2} \gamma \hat{\omega}_t^2 \hat{\sigma}_t^2 \right\} \quad (1.19)$$

dove  $\hat{\omega} = (\hat{\omega}_1, \dots, \hat{\omega}_T)$  rappresenta il vettore dei pesi stimati e  $\hat{r}_{p,t}$  è il rendimento del portafoglio al tempo  $t$ , calcolato utilizzando le stime dei pesi  $\hat{\omega}_t$ , per  $t = 1, \dots, T$  ( $\hat{r}_{p,t} = r_{f,t} + \hat{\omega}_t r_{e,t}$ ).

Per ogni istante temporale  $t$ , viene calcolato il contributo di utilità istantanea  $u_t$ . Successivamente, questi valori vengono mediati sull'intero intervallo temporale e moltiplicati per cento, al fine di ottenere una rappresentazione percentuale del contributo medio su base mensile.

#### Certainty Equivalent Return gain - $CER_{\text{gain}}$

Un criterio strettamente legato all'utilità media (1.19) è il *Guadagno di rendimento certo equivalente* rispetto alla media storica dei pesi, definito come

$$CER_{\text{gain}} = 1200 \times (CER - CER_{HA}), \quad (1.20)$$

dove

$$CER = \bar{r}_{p,t} - \frac{\gamma}{2} Var(\hat{r}_{p,t}), \quad \bar{r}_{p,t} = \frac{1}{T} \sum_{t=1}^T \hat{r}_{p,t} \quad (1.21)$$

e  $\hat{r}_{p,t} = r_{f,t} + \hat{\omega}_t r_{e,t}$ . La componente  $CER_{HA}$  si ottiene sostituendo  $\hat{\omega}_t$  con i pesi  $\bar{\omega}_t$  in 1.18, nel calcolo di  $\hat{r}_{p,t}$ . La differenza principale, rispetto al criterio dell'utilità media, risiede nell'uso della varianza campionaria del portafoglio, anziché della stima condizionata  $\hat{\sigma}_t^2$ .

Il valore  $CER_{gain}$  rappresenta il guadagno economico aggiuntivo che un investitore potrebbe ottenere accedendo alle previsioni dei pesi attraverso i predittori  $\mathbf{x}_{t-1}$ , utilizzando l'approccio del gradient boosting, rispetto al criterio di riferimento basato sulla media storica (1.18).

La differenza viene moltiplicata per 1200 per essere interpretata come una commissione percentuale annua di gestione del portafoglio. Questo indica quanto un investitore sarebbe disposto a pagare per beneficiare delle previsioni di allocazione ottimale ottenute tramite l'algoritmo di discesa del gradiente funzionale.

## Sharpe

Un altro indicatore rilevante è il *rapporto Sharpe*, calcolato dividendo il rendimento medio del portafoglio, al netto del tasso privo di rischio, per la deviazione standard dei rendimenti in eccesso:

$$Sharpe = \frac{\frac{1}{T} \sum_{t=1}^T \hat{r}_{ep,t}}{\sqrt{Var(\hat{r}_{ep,t})}} \quad (1.22)$$

dove  $\hat{r}_{ep,t} = \hat{r}_{p,t} - r_{f,t} = \hat{\omega}_t r_{e,t}$  indica il rendimento del portafoglio in eccesso rispetto al tasso risk-free.

Essendo un rapporto tra rendimento e rischio, l'indice misura quanto rendimento aggiuntivo un investitore ottiene per ogni unità di rischio assunto. Per questo motivo, si tende a preferire le strategie di allocazione che massimizzano il rapporto di Sharpe, in quanto offrono il miglior compromesso tra rendimento e rischio.



# Capitolo 2

## Dataset

Il dataset utilizzato per l'analisi dei metodi presentati include dati mensili relativi al periodo compreso tra maggio 2001 e giugno 2024. L'asset di mercato considerato è l'Euro Stoxx 50, un indice azionario introdotto nel 1998. Questo indice raggruppa le cinquanta maggiori aziende quotate in borsa nei Paesi dell'Eurozona, offrendo una rappresentazione ampia e diversificata dei principali settori industriali. In questo modo, fornisce una visione complessiva dell'andamento economico delle principali imprese operanti nell'area euro.



Figura 2.1: Serie storica dei prezzi dell'indice Euro stoxx 50

Come asset privo di rischio, si utilizza il tasso Euribor, un tasso di riferimento interbancario che riflette il costo medio a cui le principali banche europee si prestano fondi a breve termine sul mercato interbancario. L'Euribor, introdotto all'inizio del 1999 con l'avvento dell'euro, viene calcolato giornalmente su diverse scadenze ed è uno degli indicatori chiave dei mercati finanziari europei.

I rendimenti dell'indice di mercato vengono calcolati come la differenza tra i logaritmi dei prezzi in due momenti successivi, espressi dalla formula  $r_{m,t} = \log(p_t) - \log(p_{t-1})$ . I rendimenti in eccesso  $r_{e,t}$  si ottengono sottraendo al rendimento  $r_{m,t}$  dell'indice Euro Stoxx 50 il tasso di interesse Euribor, riflettendo così il rendimento netto rispetto all'asset privo di rischio.

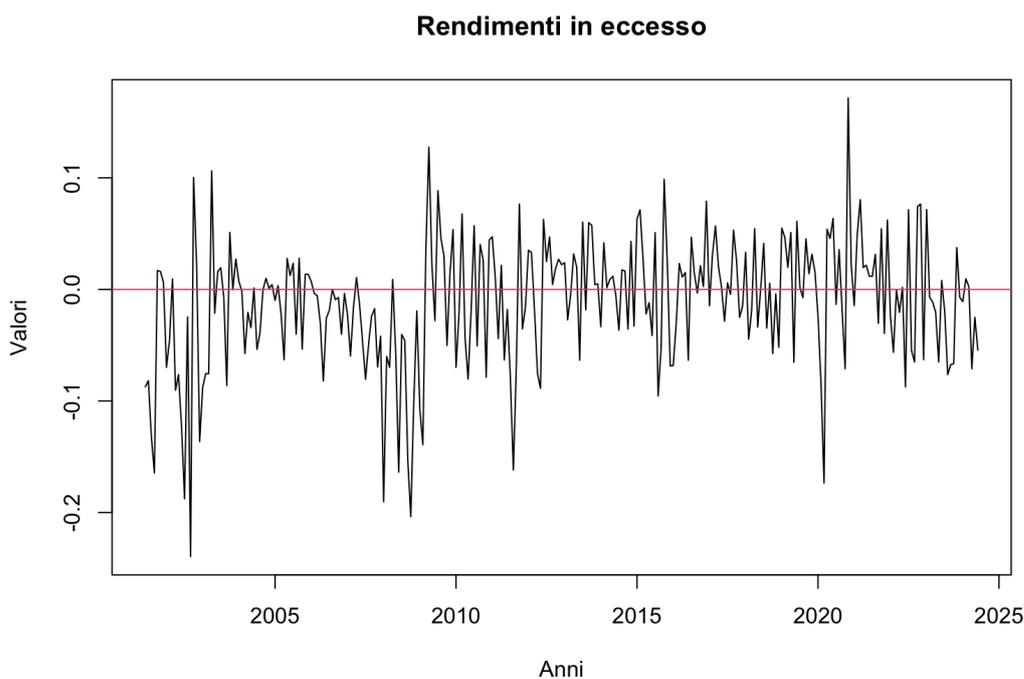


Figura 2.2: Serie storica dei rendimenti dell'indice Euro stoxx 50 in eccesso al tasso Euribor

L'insieme dei predittori contenuti in  $\mathbf{x}_{t-1}$  è composto da variabili di diversa natura, raggruppabili in due principali categorie: *Indicatori aziendali ed economici* e *Indicatori tecnici*

## 2.1 Indicatori aziendali ed economici

Il primo insieme include variabili finanziarie che forniscono una panoramica delle performance e della solidità finanziaria delle aziende che compongono l'indice Euro Stoxx 50. In questa analisi, le seguenti variabili verranno utilizzate come predittori:

- **Dividend Yield:** Misura il rapporto tra il dividendo annuale per azione e il prezzo corrente dell'azione. Esprime il rendimento percentuale che un investitore può aspettarsi sotto forma di dividendi, rispetto al prezzo pagato per l'azione e viene spesso utilizzato per valutare la capacità dell'azienda di remunerare adeguatamente il capitale investito.
- **Market value:** Il valore di mercato rappresenta il valore complessivo che il mercato attribuisce a un'azienda. Viene calcolato moltiplicando il prezzo corrente di un'azione per il numero totale di azioni in circolazione.
- **Net return:** Il rendimento netto rappresenta il rendimento complessivo di un investimento, al netto di tutte le spese associate, come commissioni di gestione, costi di transazione e imposte. Questo valore riflette il guadagno effettivo dell'investitore, fornendo una misura più accurata della performance reale dell'investimento rispetto al rendimento lordo.
- **Price to book ratio:** Il rapporto prezzo - valore contabile è un indicatore che confronta il prezzo di mercato di un titolo con il suo valore contabile. Questo indicatore viene utilizzato per valutare se un titolo è sottovalutato o sopravvalutato rispetto al valore contabile della società.
- **Price to earnings ratio:** Il rapporto prezzo - utili effettua un confronto tra prezzo di mercato di un'azione e i suoi utili. Questo indicatore fornisce un'idea di quanto gli investitori siano disposti a pagare per ogni unità di guadagno.
- **Total return index:** È un indice che misura la performance complessiva di un investimento, considerando non solo le variazioni di prezzo delle sue componenti, ma anche i dividendi e altri flussi di cassa reinvestiti.
- **Indice dei prezzi al consumo armonizzato (IPCA):** rappresenta una misura delle fluttuazioni dei prezzi nel tempo di un paniere di beni e

servizi, che riflette i consumi reali delle famiglie in un periodo specifico. In particolare questo indice viene calcolato garantendo una misurazione dell'inflazione, che consente confronti a livello europeo.

Nella tabella (2.1) vengono fornite le statistiche descrittive delle variabili descritte e dei rendimenti in eccesso al tasso Euribor in percentuale:

Tabella 2.1: Statistiche descrittive dei rendimenti in eccesso e degli Indicatori aziendali ed economici

	Media	Deviazione Standard	Minimo	Massimo	Autocorrelazione
$r_{e,t}$	0.02	5.33	-20.66	16.61	0.05
Dividend yield	2.89	0.59	1.61	5.59	0.92
Market value	2011449.91	583073.74	993129.00	3572738.00	0.97
Net return	5893.35	1938.92	2666.53	11844.45	0.97
Price/Book	1.72	0.41	0.99	3.11	0.95
Price/earnings	17.77	7.55	7.18	55.48	0.95
Total return index	1094.11	420.36	447.02	2374.34	0.98
IPCA	97.16	11.84	77.97	126.58	0.98

Il rendimento in eccesso  $r_{e,t}$  presenta un valore medio del 0.02% e una deviazione standard del 5.33%. Questi valori producono un rapporto Sharpe mensile di  $0.02\%/5.33\% = 0.0037$ , rendimento aggiuntivo ottenibile per ogni unità di rischio se tutti i pesi  $\omega_t$  fossero uguali a 1.

Si nota inoltre che tutte le variabili presentano un'elevata autocorrelazione al primo ritardo. Nell'*Appendice C* viene condotta un'analisi dei predittori per verificare la possibile presenza di distorsione nei coefficienti  $\beta$  nella regressione univariata dei rendimenti in eccesso, come indicato nella criticità 3 del metodo a due passi.

## 2.2 Indicatori tecnici

Oltre alle variabili sopra presentate, viene utilizzato un altro set di predittori dicotomici noti come indicatori tecnici.

Queste variabili sono costruite utilizzando i prezzi e i volumi storici, con l'obiettivo di identificare tendenze future nei movimenti dei prezzi e fornire segnali di acquisto o vendita.

Si analizza, quindi, la capacità di questi indicatori di prevedere i pesi e si confrontano le loro performance rispetto a quelle degli indicatori aziendali ed economici.

Le strategie utilizzate per costruire tali variabili seguono quelle proposte da Neely et al (2014) e sono le seguenti:

### 2.2.1 Indicatore Media Mobile

Il primo indicatore è un segnale basato sulla Media Mobile (MA):

$$MA(s,l) = I(MA_{s,t} \geq MA_{l,t}) = \begin{cases} 1 & \text{se } MA_{s,t} \geq MA_{l,t} \\ 0 & \text{se } MA_{s,t} < MA_{l,t} \end{cases} \quad (2.1)$$

dove  $I(\cdot)$  è la funzione indicatrice.

La media mobile  $MA_{j,t}$  viene calcolata come:

$$MA_{j,t} = \frac{1}{j} \sum_{i=0}^{j-1} P_{t-i} \quad \text{per } j = \{s,l\}, \quad s = \{1,2,3\}, \quad l = \{9,12\} \quad (2.2)$$

dove  $P_t$  rappresenta il livello di prezzo dell'indice al tempo  $t$ . I parametri  $s$  e  $l$  indicano l'ampiezza della finestra di calcolo della media mobile rispettivamente a breve e a lungo termine. Le quantità  $MA_{s,t}$  e  $MA_{l,t}$  riflettono rispettivamente le tendenze di breve e lungo periodo dei movimenti dei prezzi azionari. La media a breve termine riflette i movimenti più recenti, mentre quella a lungo termine fornisce una visione più complessiva del mercato.

Il segnale  $MA(s,l)$  viene generato dal confronto tra queste due medie. Quando il valore di media mobile a breve termine supera quella a lungo termine, la funzione indicatrice assume valore 1, fornendo un potenziale segnale di acquisto. Questo suggerisce che il mercato si trova in una fase rialzista, in cui i prezzi a breve termine stanno aumentando più rapidamente rispetto alla tendenza di lungo periodo.

Al contrario, quando la media mobile a breve termine è inferiore a quella a lungo termine, viene generato un potenziale segnale di vendita, suggerendo

una possibile fase ribassista con prezzi a breve termine in calo rispetto alla tendenza generale.

## 2.2.2 Indicatore Momentum

La seconda serie di indicatori è basata sui segnali di Momentum (MOM):

$$MOM(m) = I(P_t \geq P_{t-m}) = \begin{cases} 1 & \text{se } P_t \geq P_{t-m} \\ 0 & \text{se } P_t < P_{t-m} \end{cases} \quad m = \{9,12\} \quad (2.3)$$

Questo indicatore confronta il prezzo corrente  $P_t$  con il prezzo di  $m$  periodi precedenti  $P_{t-m}$ . Un valore positivo genera un potenziale segnale di acquisto, indicando una tendenza rialzista nei prezzi e, di conseguenza, nei rendimenti dell'indice. Al contrario, se il prezzo corrente è inferiore, si genera un segnale di vendita, suggerendo una possibile inversione di tendenza.

## 2.2.3 Indicatore sui volumi

Nelle analisi dei trend di mercato, è comune utilizzare i dati relativi ai volumi di scambio in combinazione con le serie di prezzi. Da questa considerazione nasce una terza strategia per la costruzione degli indicatori:

$$VOL(s,l) = I(MA_{s,t}^{OBV} \geq MA_{l,t}^{OBV}) = \begin{cases} 1 & \text{se } MA_{s,t}^{OBV} \geq MA_{l,t}^{OBV} \\ 0 & \text{se } MA_{s,t}^{OBV} < MA_{l,t}^{OBV} \end{cases} \quad (2.4)$$

In questa espressione,  $MA_{s,t}^{OBV}$  è calcolata attraverso la media mobile dell'*On-Balance Volume (OBV)* definita come:

$$MA_{j,t}^{OBV} = \frac{1}{j} \sum_{i=0}^{j-1} OBV_{t-i} \quad \text{per } j = \{s,l\}, \quad s = \{1,2,3\}, \quad l = \{9,12\} \quad (2.5)$$

L'OBV è calcolato come la somma cumulativa dei volumi, pesati in base alla direzione del prezzo, secondo la seguente formula:

$$OBV_t = \sum_{k=1}^t VOL_k D_k \quad (2.6)$$

dove  $VOL_k$  rappresenta il volume scambiato durante il periodo  $k$  e  $D_k$  è una variabile dicotomica così definita:

$$D_k = 2I(P_k - P_{k-1}) - 1 = \begin{cases} 1 & \text{se } P_k - P_{k-1} \geq 0 \\ -1 & \text{se } P_k - P_{k-1} < 0 \end{cases} \quad (2.7)$$

Un elevato volume di scambi, unito a recenti aumenti dei prezzi, suggerisce un trend di mercato positivo; in questo contesto  $VOL(s,1)$  produce un segnale di acquisto.



# Capitolo 3

## Risultati

In questa sezione verranno riportati ed analizzati i risultati relativi alle performance dei pesi di portafoglio  $\omega_t$ , ottenuti con i metodi di discesa del gradiente e di regressione lineare predittiva esposti nel *Capitolo 1*. Per entrambi i metodi viene utilizzato lo stesso set di predittori descritto nel *Capitolo 2*.

Essendo che l'obiettivo è di previsione, l'analisi è stata effettuata secondo una procedura *Out of sample*, ovvero, i pesi di portafoglio per il mese  $t$  vengono costruiti utilizzando le informazioni disponibili al tempo  $t - 1$ , contenute in  $\mathbf{x}_{t-1}$ , insieme ai valori stimati di varianza  $\sigma_t^2$ .

I due metodi utilizzano una finestra di stima mobile come dati di addestramento, e le performance delle previsioni di portafoglio finali vengono valutate in termini di utilità media, guadagno di rendimento certo equivalente ( $CER_{gain}$ ) e rapporto di Sharpe, come definiti nel *Capitolo 1*.

L'obiettivo principale dell'analisi è confrontare il comportamento del nuovo metodo a *1 passo* con il metodo precedente basato sulla regressione lineare, su un ampio insieme di predittori, valutando quali variabili possiedono maggiore potere predittivo nella determinazione dei pesi e quali possiedono informazioni utili alla previsione a confronto con i criteri di riferimento, benchmark, i quali si riferiscono a una previsione dei pesi senza informazione.

È importante sottolineare che il metodo di utility boosting incorpora un elemento di contrazione (il parametro  $\nu$ ), che offre una protezione contro potenziali problemi di sovradattamento. Di conseguenza, considerando anche che i due metodi non sono annidati, i criteri di valutazione del metodo a *1 passo* non risulteranno automaticamente superiori a quelli del metodo a *2 passi*.

Viene condotta un'analisi sia a livello univariato, per valutare le performance di previsione dei singoli predittori, sia a livello multivariato, utilizzando gruppi di variabili contemporaneamente: indicatori economici e aziendali (ECO/AZ), indicatori tecnici (TECH) e l'intero insieme di predittori (ALL). Uno dei vantaggi del metodo basato sulla discesa del gradiente è l'assenza di necessità di preselezionare le variabili, poiché, grazie all'uso di un modello di base ad albero, la selezione viene effettuata automaticamente. Per catturare le interazioni tra le variabili, si utilizza un albero a due livelli. Al contrario, l'approccio a *2 passi* richiede un'ulteriore fase di selezione, in quanto il modello lineare non effettua tale scelta. In particolare, poiché le variabili esplicative presentano multicollinearità, l'applicazione diretta del metodo dei minimi quadrati non è possibile; pertanto, si utilizza l'analisi delle componenti principali (PCA) per incorporare la maggior parte delle informazioni dalle variabili, riducendo il rischio di sovradattamento e risolvendo il problema della multicollinearità.

Per verificare eventuali differenze nei risultati, vengono testati diversi valori per i parametri  $\omega^{max}$  e  $\gamma$ . In particolare, per il parametro  $\gamma$  di avversione al rischio, si utilizzano i valori 5 e 3, al fine di rappresentare rispettivamente un investitore con elevata avversione al rischio ed uno con bassa avversione.

### 3.1 Limite superiore $\omega^{max} = 1$

È stata condotta un'analisi preliminare sulla dimensione della finestra di stima per comprendere come i risultati cambino quando ai modelli vengono forniti diversi quantitativi di dati per l'addestramento. Sono state utilizzate due finestre temporali: una di 5 anni (60 osservazioni) e una di 10 anni (120 osservazioni). Viene esaminato il caso in cui  $\gamma = 5$ , e  $\omega^{max} = 1$ . I risultati riportati mostrano come la seconda opzione sia preferibile:

Tabella 3.1: Risultati dell'analisi univariata con  $\gamma = 5$  e  $\omega^{max} = 1$  e una finestra = 60

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
Indicatori Aziendali ed Economici						
Dividend Yield	0.414	1.864	0.098	0.249	0.225	0.047
Market Value	0.439	2.127	0.099	0.414	1.944	0.116
Net return	0.335	<b>0.502</b>	0.039	0.421	1.996	0.118
Price-Book ratio	0.448	1.910	0.094	0.387	1.625	0.084
Price-Earnings ratio	<b>0.198</b>	<b>-0.558</b>	<b>0.012</b>	0.334	1.917	0.093
Total return index	0.505	2.909	0.132	0.340	1.092	0.049
IPCA	<b>-0.016</b>	<b>-2.910</b>	<b>-0.047</b>	<b>0.040</b>	<b>-2.375</b>	<b>-0.092</b>
Indicatori Tecnici						
MA(1,9)	<b>0.199</b>	<b>-0.307</b>	<b>-0.016</b>	<b>0.172</b>	<b>-1.064</b>	<b>-0.025</b>
MA(1,12)	<b>0.191</b>	<b>-0.738</b>	<b>-0.028</b>	<b>0.198</b>	<b>-0.821</b>	<b>-0.015</b>
MA(2,9)	<b>0.274</b>	0.745	0.024	<b>0.143</b>	<b>-0.891</b>	<b>-0.047</b>
MA(2,12)	<b>0.265</b>	<b>0.550</b>	<b>0.013</b>	<b>0.135</b>	<b>-1.035</b>	<b>-0.049</b>
MA(3,9)	<b>0.194</b>	<b>-0.269</b>	<b>-0.026</b>	<b>0.097</b>	<b>-1.660</b>	<b>-0.070</b>
MA(3,12)	<b>0.203</b>	<b>-0.290</b>	<b>-0.006</b>	<b>0.192</b>	<b>-0.533</b>	<b>-0.011</b>
MOM(9)	0.345	1.244	0.069	0.260	0.165	0.025
MOM(12)	0.423	1.946	0.098	0.404	1.746	0.091
VOL(1,9)	0.338	0.881	0.058	0.264	0.234	0.033
VOL(1,12)	0.291	<b>0.338</b>	0.038	0.293	0.566	0.047
VOL(2,9)	<b>0.130</b>	<b>-0.998</b>	<b>-0.051</b>	<b>0.185</b>	<b>-0.529</b>	-0.007
VOL(2,12)	0.288	0.871	0.049	0.220	<b>-0.315</b>	0.016
VOL(3,9)	<b>0.264</b>	0.826	0.035	<b>0.200</b>	<b>-0.061</b>	0.010
VOL(3,12)	<b>0.251</b>	<b>0.407</b>	0.028	0.220	0.020	0.018
Benchmark						
	0.276	0.725	0.021	0.218	0.000	-0.011

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Tabella 3.2: Risultati dell'analisi univariata con  $\gamma = 5$  e  $\omega^{max} = 1$  e una finestra = 120

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
Indicatori Aziendali ed Economici						
Dividend Yield	0.306	0.881	0.032	0.323	1.092	0.054
Market Value	0.367	1.558	0.058	0.304	1.033	0.025
Net return	0.353	1.362	0.045	0.303	1.017	-0.032
Price-Book ratio	0.377	1.649	0.063	0.336	1.416	0.056
Price-Earnings ratio	0.258	0.657	0.021	0.243	0.651	0.020
Total return index	0.467	2.861	0.151	0.283	0.780	-0.061
IPCA	<b>0.196</b>	<b>-0.075</b>	0.010	<b>0.195</b>	<b>-0.027</b>	0.017
Indicatori Tecnici						
MA(1,9)	<b>0.206</b>	<b>0.054</b>	-0.049	0.237	0.026	-0.039
MA(1,12)	0.256	0.630	-0.032	0.227	0.015	-0.041
MA(2,9)	<b>0.231</b>	<b>0.365</b>	<b>-0.060</b>	<b>0.168</b>	<b>-0.409</b>	<b>-0.134</b>
MA(2,12)	0.241	0.483	-0.049	<b>0.197</b>	<b>-0.141</b>	<b>-0.094</b>
MA(3,9)	0.240	0.448	-0.051	0.243	0.435	-0.020
MA(3,12)	<b>0.232</b>	<b>0.372</b>	<b>-0.058</b>	<b>0.168</b>	<b>-0.673</b>	<b>-0.093</b>
MOM(9)	0.293	1.021	0.011	0.299	0.823	0.016
MOM(12)	0.346	1.382	0.059	0.364	1.504	0.065
VOL(1,9)	0.254	0.458	-0.012	0.209	0.084	-0.041
VOL(1,12)	<b>0.223</b>	<b>0.113</b>	-0.044	<b>0.187</b>	<b>-0.585</b>	-0.053
VOL(2,9)	<b>0.166</b>	<b>-0.495</b>	<b>-0.106</b>	<b>0.113</b>	<b>-1.438</b>	<b>-0.100</b>
VOL(2,12)	<b>0.161</b>	<b>-0.681</b>	<b>-0.081</b>	<b>0.127</b>	<b>-1.236</b>	<b>-0.098</b>
VOL(3,9)	<b>0.183</b>	<b>-0.370</b>	<b>-0.069</b>	<b>0.105</b>	<b>-1.471</b>	<b>-0.096</b>
VOL(3,12)	<b>0.179</b>	<b>-0.409</b>	<b>-0.055</b>	<b>0.114</b>	<b>-1.160</b>	<b>-0.089</b>
Benchmark						
	0.237	0.446	-0.053	0.208	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Tabella 3.3: Risultati dell'analisi multivariata con  $\gamma = 5$  e  $\omega^{max} = 1$  e finestra = 60

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
ECO/AZ	<b>0.263</b>	<b>0.636</b>	0.079	0.316	1.057	0.123
TECH	0.431	2.338	0.112	<b>0.148</b>	<b>-1.416</b>	<b>-0.017</b>
ALL	0.303	0.800	0.046	0.305	0.927	0.053
Benchmark						
	0.276	0.725	0.021	0.218	0.000	-0.011

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Tabella 3.4: Risultati dell'analisi multivariata con  $\gamma = 5$  e  $\omega^{max} = 1$  e finestra = 120

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
ECO/AZ	0.456	1.894	0.159	0.359	1.541	0.152
TECH	0.239	0.450	0.005	<b>0.104</b>	<b>-1.770</b>	-0.053
ALL	0.412	1.801	0.143	0.336	1.394	0.144
Benchmark						
	0.237	0.446	-0.053	0.208	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Aumentando la dimensione della finestra di stima emergono alcune differenze rilevanti: innanzitutto, i valori dei criteri di riferimento, riportati alla base delle tabelle, si riducono, suggerendo che l'aumento delle informazioni passate riduce l'importanza della "non informazione" nel potere predittivo.

In tutte le tabelle sono evidenziati in grassetto i criteri di valutazione inferiori ai benchmark; le variabili corrispondenti a questi criteri non mostrano potere predittivo poichè la previsione senza informazione rappresentata dai criteri di riferimento (benchmark) presenta valori superiori. In particolare, per il metodo di utility boosting a 1 passo, una finestra di stima di 10 anni (120 osservazioni) consente a un numero maggiore di variabili di dimostrare potere predittivo rispetto ai criteri di riferimento, rispetto a una finestra di stima di 5 anni. Inoltre, l'aumento delle informazioni rende i criteri di valutazione più omogenei nel determinare se le variabili esplicative possiedono o meno potere predittivo nella determinazione dei pesi di portafoglio. Questo effetto non si osserva con una finestra di stima più ridotta, in cui i diversi criteri portano a preferenze meno allineate.

Oltre a tutto questo, con una finestra di 120 osservazioni, i criteri di valutazione si allineano maggiormente all'interno dello stesso modello nella selezione delle variabili con maggiore potere predittivo. Nel caso univariato, tra le variabili aziendali ed economiche, assumono particolare rilevanza il *Total Return Index* e il *Price to Book Ratio* nel metodo di utility boosting, mentre nel metodo basato sulla regressione lineare si distinguono il *Price to Book Ratio* e il *Dividend Yield*, anche se con valori inferiori rispetto alle migliori variabili del primo metodo.

Per quanto riguarda gli indicatori tecnici, nel metodo a un passo MOM(12), MOM(9) e MA(1,12) mostrano le migliori prestazioni. Il metodo lineare concorda sui momentum, ma considera migliore la media mobile MA(3,9) rispetto a MA(1,12). Entrambi i metodi considerano senza potere predittivo utile le variabili IPCA, MA(2,9), MA(3,12) e tutti i volumi, eccetto VOL(1,9) dato che i relativi risultati sono inferiori ai valori benchmark; tuttavia, differiscono nella valutazione delle variabili MA(1,9) e MA(2,12).

Nel caso multivariato, per entrambi i metodi, i valori più alti dei vari criteri si riscontrano nel gruppo delle variabili aziendali ed economiche. Gli indicatori tecnici mostrano un potere predittivo limitato nel primo modello, mentre nel secondo non ne evidenziano affatto, poichè i loro valori risultano inferiori ai

benchmark.

Poiché una finestra di stima più ampia produce risultati più omogenei e rilevanti, essa sarà impiegata anche nelle analisi successive con diversi valori di  $\gamma$  e  $\omega^{max}$ .

Tabella 3.5: Risultati dell'analisi univariata con  $\gamma = 3$  e  $\omega^{max} = 1$

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
	Indicatori Aziendali ed Economici					
Dividend Yield	0.432	3.183	0.078	0.376	2.481	0.075
Market Value	0.423	2.993	0.074	0.311	1.800	-0.021
Net return	0.396	2.645	0.058	0.296	1.774	-0.032
Price-Book ratio	0.424	3.090	0.076	0.352	2.442	0.058
Price-Earnings ratio	0.196	0.539	-0.006	0.183	0.729	-0.022
Total return index	0.511	4.236	0.150	0.292	1.420	-0.057
IPCA	<b>0.082</b>	<b>-0.717</b>	-0.026	0.301	1.876	0.041
	Indicatori Tecnici					
MA(1,9)	0.144	0.114	-0.078	0.250	1.085	-0.012
MA(1,12)	0.138	0.063	-0.073	0.206	0.445	-0.031
MA(2,9)	<b>0.084</b>	<b>-0.548</b>	<b>-0.108</b>	<b>0.072</b>	<b>-0.685</b>	<b>-0.134</b>
MA(2,12)	<b>0.095</b>	<b>-0.414</b>	<b>-0.100</b>	<b>0.124</b>	<b>-0.144</b>	<b>-0.095</b>
MA(3,9)	0.105	-0.306	-0.093	0.224	0.918	-0.030
MA(3,12)	<b>0.087</b>	<b>-0.522</b>	<b>-0.104</b>	<b>0.114</b>	<b>-0.440</b>	-0.083
MOM(9)	0.166	0.348	-0.058	0.329	2.053	0.034
MOM(12)	0.391	2.731	0.062	0.462	3.533	0.089
VOL(1,9)	0.171	0.385	-0.049	0.215	0.665	-0.026
VOL(1,12)	0.150	0.143	-0.066	0.176	0.107	-0.042
VOL(2,9)	<b>0.042</b>	<b>-1.093</b>	<b>-0.116</b>	<b>0.051</b>	<b>-1.370</b>	<b>-0.089</b>
VOL(2,12)	<b>0.038</b>	<b>-1.398</b>	<b>-0.097</b>	<b>0.046</b>	<b>-1.397</b>	<b>-0.091</b>
VOL(3,9)	<b>0.033</b>	<b>-1.508</b>	<b>-0.098</b>	<b>0.024</b>	<b>-1.549</b>	<b>-0.095</b>
VOL(3,12)	<b>0.056</b>	<b>-1.041</b>	-0.078	<b>0.036</b>	<b>-1.264</b>	-0.081
	Benchmark					
	0.104	-0.311	-0.095	0.138	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Tabella 3.6: Risultati dell'analisi multivariata con  $\gamma = 3$  e  $\omega^{max} = 1$

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
ECO/AZ	0.584	2.393	0.192	0.492	2.128	0.146
TECH	0.161	0.160	-0.023	<b>0.117</b>	<b>-0.571</b>	-0.049
ALL	0.539	2.053	0.175	0.448	1.905	0.130
	Benchmark					
	0.104	-0.311	-0.095	0.138	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Sono evidenziati in grassetto i criteri di valutazione inferiori rispetto ai valori di riferimento riportati alla base delle tabelle. Anche in questo caso, nell'analisi univariata, i due metodi concordano nel considerare prive di potere predittivo le stesse variabili esplicative: MA(2,9), MA(2,12), MA(3,12), VOL(2,9), VOL(2,12), VOL(3,9) e VOL(3,12). Tuttavia, differiscono nella valutazione della variabile *IPCA*, ritenuta non rilevante dal primo metodo ma informativa dal secondo.

Tra le variabili economiche e aziendali, il primo metodo attribuisce maggiore potere predittivo al *Total return Index* e al *Dividend Yield*, mentre il metodo a 2 passi riconferma le variabili *Dividend Yield* e *Price to Book ratio* come più informative per la previsione, come nel caso con  $\gamma = 5$ .

Per quanto riguarda gli indicatori tecnici, nel metodo a un passo MOM(12), VOL(1,9), MOM(9) e VOL(1,12) mostrano le migliori prestazioni. Il metodo lineare concorda sui momentum, ma considera migliori le medie mobili MA(1,9) e MA(3,9) rispetto agli indicatori sui volumi.

Nell'analisi multivariata si confermano i risultati ottenuti con  $\gamma = 5$ , mostrando un maggiore potere predittivo per le variabili economiche e aziendali, e una rilevanza per la previsione limitata o nulla per gli indicatori tecnici.

In entrambe le analisi si osserva che le variabili economiche e aziendali presentano valori più elevati rispetto agli indicatori tecnici, sia nel caso univariato che in quello multivariato. In particolare, per  $\gamma = 3$ , questi valori risultano più alti poiché l'investitore mostra una maggiore propensione al rischio.

Il troncamento dei pesi nel metodo a *2 passi*, per soddisfare i limiti (1.9), si nota particolarmente nei grafici di alcune variabili riportate nell'*Appendice D*, dove sono presenti diverse sezioni piatte, specialmente al livello del limite inferiore  $\omega^{min} = 0$ , mentre l'approccio di utility boosting riesce a mantenere automaticamente i valori all'interno dell'intervallo.

## 3.2 Limite superiore $\omega^{max} = 1.5$

Viene analizzato un caso aggiuntivo in cui si assume che il limite superiore  $\omega^{max}$  in (1.9) sia pari a 1.5. Questo valore implica l'uso della leva finanziaria, permettendo all'investitore di allocare fino al 150% del capitale nel mercato: il 100% di capitale proprio e un ulteriore 50% ottenuto a prestito.

I risultati riportati di seguito mostrano alcune differenze rispetto ai casi precedenti in cui  $\omega^{max} = 1$ .

Tabella 3.7: Risultati dell'analisi univariata con  $\gamma = 5$  e  $\omega^{max} = 1.5$

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
	Indicatori Aziendali ed Economici					
Dividend Yield	0.198	<b>-0.320</b>	0.002	0.309	1.384	0.052
Market Value	0.363	1.571	0.058	0.304	1.033	0.025
Net return	0.361	1.525	0.056	0.303	1.017	-0.032
Price-Book ratio	0.299	0.539	0.009	0.336	1.416	0.056
Price-Earnings ratio	0.210	0.475	0.027	0.234	0.982	0.060
Total return index	0.448	2.530	0.123	0.283	0.780	-0.061
IPCA	<b>0.035</b>	<b>-1.403</b>	-0.032	<b>0.108</b>	<b>-0.767</b>	0.010
	Indicatori Tecnici					
MA(1,9)	<b>0.156</b>	<b>-0.614</b>	-0.060	0.213	<b>-0.314</b>	-0.029
MA(1,12)	0.221	0.242	-0.051	0.227	0.015	-0.041
MA(2,9)	<b>0.197</b>	<b>0.002</b>	<b>-0.076</b>	<b>0.168</b>	<b>-0.409</b>	<b>-0.134</b>
MA(2,12)	0.203	0.065	-0.070	<b>0.197</b>	<b>-0.141</b>	<b>-0.094</b>
MA(3,9)	<b>0.197</b>	<b>0.005</b>	<b>-0.076</b>	0.243	0.435	-0.020
MA(3,12)	0.199	0.021	-0.074	<b>0.168</b>	<b>-0.673</b>	<b>-0.093</b>
MOM(9)	0.252	0.558	-0.024	0.297	0.801	0.015
MOM(12)	0.282	0.540	0.038	0.336	1.056	0.052
VOL(1,9)	<b>0.188</b>	<b>-0.418</b>	-0.047	<b>0.195</b>	<b>-0.459</b>	-0.047
VOL(1,12)	<b>0.132</b>	<b>-1.027</b>	-0.075	<b>0.166</b>	<b>-0.916</b>	-0.061
VOL(2,9)	<b>0.134</b>	<b>-0.836</b>	<b>-0.116</b>	<b>0.113</b>	<b>-1.438</b>	<b>-0.100</b>
VOL(2,12)	<b>0.113</b>	<b>-1.147</b>	<b>-0.115</b>	<b>0.127</b>	<b>-1.236</b>	<b>-0.098</b>
VOL(3,9)	<b>0.148</b>	<b>-0.747</b>	<b>-0.088</b>	<b>0.105</b>	<b>-1.471</b>	<b>-0.096</b>
VOL(3,12)	<b>0.118</b>	<b>-0.958</b>	<b>-0.081</b>	<b>0.114</b>	<b>-1.160</b>	<b>-0.089</b>
	Benchmark					
	0.197	0.003	-0.076	0.208	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Tabella 3.8: Risultati dell'analisi multivariata con  $\gamma = 5$  e  $\omega^{max} = 1.5$

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
ECO/AZ	0.517	2.015	0.172	0.495	1.928	0.174
TECH	<b>0.019</b>	<b>-3.683</b>	-0.047	<b>0.043</b>	<b>-2.581</b>	-0.046
ALL	0.451	1.322	0.162	0.441	1.312	0.134
	Benchmark					
	0.197	0.003	-0.076	0.208	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Complessivamente, rispetto al caso con  $\omega^{max} = 1$ , i criteri mostrano valori inferiori, in particolare nel metodo non parametrico. Nel caso lineare, molti criteri risultano immutati, indicando che la troncatura dei pesi (sia con  $\omega^{max} = 1$  sia con  $\omega^{max} = 1.5$ ) non ha influenzato la loro stima. Questa stabilità emerge anche dai grafici in *Appendice E* e dai valori dei benchmark, che rimangono invariati nei due scenari con diversi limiti superiori.

Le principali differenze emergono nel metodo a *1 passo*. Questo metodo non considera informativi per la previsione tutti gli indicatori di volume (come nel metodo lineare) e l'insieme degli indicatori tecnici (TECH), nonostante i valori dei criteri di riferimento siano diminuiti. Inoltre, mostra preferenze differenti per gli indicatori di media mobile MA(3,9) e MA(3,12).

Tabella 3.9: Risultati dell'analisi univariata con  $\gamma = 3$  e  $\omega^{max} = 1.5$

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
Indicatori Aziendali ed Economici						
Dividend Yield	0.565	4.761	0.119	0.346	2.061	0.057
Market Value	0.426	2.768	0.062	0.298	1.800	0.021
Net return	0.358	2.243	0.038	0.296	1.774	-0.032
Price-Book ratio	0.354	1.599	0.026	0.352	2.442	0.058
Price-Earnings ratio	<b>0.097</b>	<b>-0.453</b>	-0.023	0.191	1.229	0.014
Total return index	0.550	4.460	0.136	0.263	1.384	-0.061
IPCA	<b>0.093</b>	<b>-0.463</b>	0.004	0.146	0.192	0.017
Indicatori Tecnici						
MA(1,9)	<b>0.154</b>	<b>0.271</b>	-0.038	0.202	0.295	-0.017
MA(1,12)	0.197	0.813	-0.032	0.175	0.021	-0.040
MA(2,9)	<b>0.143</b>	<b>0.207</b>	<b>-0.064</b>	<b>0.072</b>	<b>-0.685</b>	<b>-0.134</b>
MA(2,12)	0.180	0.624	-0.044	<b>0.119</b>	<b>-0.204</b>	<b>-0.094</b>
MA(3,9)	0.159	0.402	<b>-0.055</b>	0.197	0.747	-0.039
MA(3,12)	<b>0.146</b>	<b>0.245</b>	<b>-0.062</b>	<b>0.078</b>	<b>-0.938</b>	<b>-0.092</b>
MOM(9)	0.255	1.462	0.000	0.300	1.654	0.021
MOM(12)	0.346	2.184	0.050	0.415	2.830	0.070
VOL(1,9)	0.212	0.844	-0.012	0.151	<b>-0.127</b>	-0.038
VOL(1,12)	<b>0.139</b>	<b>0.037</b>	-0.051	<b>0.115</b>	<b>-0.667</b>	-0.051
VOL(2,9)	<b>0.063</b>	<b>-0.805</b>	<b>-0.098</b>	<b>-0.013</b>	<b>-2.242</b>	<b>-0.099</b>
VOL(2,12)	<b>0.067</b>	<b>-1.013</b>	<b>-0.073</b>	<b>0.003</b>	<b>-2.013</b>	<b>-0.098</b>
VOL(3,9)	<b>0.065</b>	<b>-1.003</b>	<b>-0.072</b>	<b>-0.026</b>	<b>-2.346</b>	<b>-0.094</b>
VOL(3,12)	<b>-0.101</b>	<b>-2.685</b>	<b>-0.128</b>	<b>-0.019</b>	<b>-2.004</b>	<b>-0.089</b>
Benchmark						
	0.158	0.398	-0.055	0.138	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Tabella 3.10: Risultati dell'analisi multivariata con  $\gamma = 3$  e  $\omega^{max} = 1.5$

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
ECO/AZ	0.450	1.973	0.148	0.419	1.876	0.151
TECH	<b>-0.015</b>	<b>-1.657</b>	<b>-0.069</b>	<b>0.005</b>	<b>-2.233</b>	-0.052
ALL	0.406	1.718	0.130	0.403	1.691	0.136
	Benchmark					
	0.158	0.398	-0.055	0.138	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Nel complesso, in questo secondo caso, con il metodo di boosting univariato si ottengono valori più elevati rispetto al caso con  $\omega^{max} = 1$ , mentre con il metodo di regressione lineare i valori si riducono, così come il numero di variabili con valori immutati. Ciò suggerisce che, con una maggiore avversione al rischio, la troncatura del peso, al limite superiore  $\omega = 1.5$ , abbia un'influenza maggiore. Anche in questo caso, i criteri di riferimento nella regressione lineare restano invariati, mentre nel metodo a *1 passo* si osserva un incremento.

Nel metodo non parametrico, inoltre, emergono differenze nel potere predittivo delle variabili, specialmente per gli indicatori tecnici: i due casi, con diversi limiti superiori, mostrano preferenze differenti per le variabili MA(1,9), MA(2,12), VOL(1,12) e per l'insieme degli indicatori tecnici (TECH) rispetto ai criteri di riferimento: in questo caso,  $\omega^{max} = 1.5$ , tutte queste valutazioni, ad eccezione di MA(2,12) risultano inferiori ai benchmark, mentre MA(2,12) acquista rilevanza poichè supera i valori dei criteri di riferimento. Tra le variabili aziendali ed economiche, la variabile *Price to Earnings ratio* perde la sua rilevanza rispetto ai benchmark, non mostrando più potere predittivo utile.

Variando i parametri di avversione al rischio  $\gamma$  e il limite superiore  $\omega^{max}$ , si ottengono diverse valutazioni sia a livello univariato che multivariato. Tuttavia, i metodi giungono sempre alle stesse conclusioni, indicando l'insieme delle variabili economiche e aziendali (ECO/AZ) come quello con maggiore potere predittivo rispetto agli indicatori tecnici. Anche nelle analisi univariate, queste variabili mantengono una capacità predittiva costantemente superiore.

Al contrario, gli indicatori tecnici, con l'eccezione di quelli di momentum, mostrano generalmente una rilevanza limitata, con valori che risultano inferiori o solo leggermente superiori rispetto ai benchmark. Anche nell'analisi multivariata, non evidenziano mai un potere predittivo rilevante.

In generale si può notare che l'approccio non parametrico restituisce risultati più elevati nella valutazione delle variabili economiche e aziendali, sia nel caso univariato che in quello multivariato, rispetto al modello lineare. Per quanto riguarda gli indicatori tecnici, invece, i risultati non mostrano una chiara predominanza di un modello sull'altro. Questo è probabilmente dovuto al fatto che molte di queste variabili non sembrano offrire un potere predittivo rilevante, né singolarmente né in combinazione.

### 3.3 Differenze prime dei predittori economici ed aziendali

Le analisi riportate in *Appendice C* mostrano che le variabili economiche e aziendali presentano un'elevata persistenza, e il test di Dickey-Fuller aumentato (ADF) conferma la presenza di radice unitaria in queste variabili. Si esamina quindi l'impatto dell'uso delle differenze prime di tali predittori sulle previsioni fuori campione, sia in contesti univariati che multivariati, dove solo le variabili di interesse sono differenziate. Vengono analizzati i casi in cui il limite superiore  $\omega^{max}$  è pari a 1, per diversi valori di  $\gamma$ .

I risultati sono riportati nelle tabelle seguenti:

Tabella 3.11: Analisi con parametri  $\gamma = 5$  e  $\omega^{max} = 1$  con differenze prime delle variabili persistenti

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
Indicatori Aziendali ed Economici						
$\Delta$ Dividend Yield	0.243	0.538	-0.023	0.251	0.546	-0.010
$\Delta$ Market Value	<b>0.137</b>	<b>-0.674</b>	-0.047	<b>0.195</b>	<b>-0.106</b>	-0.034
$\Delta$ Net Return	<b>0.207</b>	<b>0.172</b>	-0.019	<b>0.192</b>	<b>-0.383</b>	-0.041
$\Delta$ Price-Book Ratio	<b>0.194</b>	<b>-0.423</b>	-0.027	<b>0.130</b>	<b>-0.826</b>	<b>-0.097</b>
$\Delta$ Price-Earnings ratio	<b>0.134</b>	<b>-1.221</b>	<b>-0.070</b>	<b>0.141</b>	<b>-0.931</b>	<b>-0.096</b>
$\Delta$ Total Return Index	<b>0.223</b>	<b>0.270</b>	-0.007	<b>0.184</b>	<b>-0.546</b>	-0.042
$\Delta$ IPCA	<b>0.032</b>	<b>-2.086</b>	<b>-0.102</b>	<b>0.175</b>	<b>-0.271</b>	-0.078
Multivariato						
ECO/AZ	<b>0.108</b>	<b>-0.280</b>	<b>-0.089</b>	<b>0.098</b>	<b>-0.631</b>	-0.046
ALL	<b>0.097</b>	<b>-0.476</b>	<b>-0.100</b>	<b>0.084</b>	<b>-1.507</b>	-0.055
Benchmark						
	0.230	0.449	-0.053	0.200	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

I criteri di valutazione rispetto ai risultati nelle tabelle (3.2) e (3.4) si riducono notevolmente, e molte variabili, i cui criteri di valutazione sono evidenziati in grassetto, non mostrano un potere predittivo significativo rispetto ai criteri

di riferimento, poichè i loro valori risultano inferiori ai benchmark, sia nel caso univariato che in quello multivariato.

Tabella 3.12: Analisi con parametri  $\gamma = 3$  e  $\omega^{max} = 1$  con differenze prime delle variabili persistenti

Variabili	Utility Boosting			Lineare		
	Utilità (%)	Cer (%)	Sharpe	Utilità (%)	Cer (%)	Sharpe
Indicatori Aziendali ed Economici						
$\Delta$ Dividend Yield	0.190	0.603	-0.044	0.183	0.565	-0.032
$\Delta$ Market Value	0.197	0.932	-0.011	0.143	0.126	-0.044
$\Delta$ Net Return	0.263	1.682	0.016	<b>0.128</b>	<b>-0.405</b>	-0.050
$\Delta$ Price-Book Ratio	0.237	1.385	0.006	<b>0.056</b>	<b>-0.909</b>	<b>-0.089</b>
$\Delta$ Price-Earnings ratio	0.197	0.478	-0.034	<b>0.045</b>	<b>-1.322</b>	<b>-0.100</b>
$\Delta$ Total Return Index	0.210	1.021	-0.010	<b>0.129</b>	<b>-0.425</b>	-0.049
$\Delta$ IPCA	<b>0.037</b>	<b>-1.266</b>	-0.069	<b>0.099</b>	<b>-0.293</b>	-0.081
Multivariato						
ECO/AZ	0.133	0.140	-0.049	<b>0.118</b>	<b>-0.530</b>	-0.048
ALL	0.130	0.136	-0.051	<b>0.115</b>	<b>-0.599</b>	-0.053
Benchmark						
	0.096	-0.313	-0.096	0.130	0.000	-0.086

I valori dei criteri di valutazione inferiori ai benchmark, riportati alla base della tabella, sono evidenziati in grassetto; le variabili corrispondenti a questi criteri non mostrano potere predittivo.

Anche in questo secondo caso, i valori, rispetto a quelli riportati nelle tabelle (3.5) e (3.6), si sono ridotti considerevolmente. Tuttavia, nel metodo non parametrico univariato, il numero di variabili con valori maggiori rispetto ai benchmark aumenta rispetto al caso precedente con un diverso valore  $\gamma$ . Nell'analisi multivariata, il metodo non parametrico dimostra una certa capacità predittiva rispetto ai benchmark, sebbene non sia molto elevata. Al contrario, i valori del modello basato sulla regressione lineare non riescono a superare i criteri di riferimento, confermando, come nel caso precedente, che le combinazioni di variabili non apportano potere predittivo utile rispetto alla previsione senza informazione rappresentata dai benchmark.

In entrambi i casi, il metodo a *1 passo* mantiene la sua predominanza rispetto al modello lineare, mostrando valori dei criteri di valutazione generalmente più elevati e in generale maggiori rispetto ai criteri di riferimento.

Tuttavia, si osserva che il rapporto *Sharpe* è spesso negativo, evidenziando come il rischio assunto non sia adeguatamente compensato, a differenza dei risultati ottenuti dalle analisi con le variabili economiche e aziendali non in differenza prima, dove i valori del *Sharpe* erano generalmente positivi. Analoghe conclusioni emergono dal criterio *CER*: nelle analisi non in differenza, i valori erano generalmente positivi, mentre utilizzando le differenze, i valori risultano spesso negativi, in particolare nel modello parametrico a *2 passi*.

Nonostante le variabili analizzate siano altamente persistenti e la presenza di radice unitaria sia confermata, l'utilizzo delle differenze prime non sembra migliorare le capacità predittive dei modelli. Infatti, i risultati delle analisi mostrano che i criteri di valutazione evidenziano una perdita di potere predittivo rispetto ai valori non in differenza.



# Conclusione

Questo lavoro propone un metodo non parametrico e flessibile per l'allocazione ottimale del portafoglio, in contrasto con gli approcci classici a due fasi comunemente adottati. Invece di ricorrere a stime intermedie dei rendimenti, questo metodo collega direttamente le variabili predittive ai pesi di portafoglio massimizzando la funzione di utilità sottostante, rappresentativa dell'investitore. Ciò permette di evitare i limiti della prevedibilità statistica dei rendimenti, che non rappresenta l'interesse principale e che potrebbe introdurre errori di specificazione e rumore nelle stime dei pesi.

Si svolge l'analisi in un contesto semplice in cui la scelta d'investimento è tra un tasso privo di rischio e l'indice azionario *Euro Stoxx 50*, asset rischioso.

L'approccio di boosting personalizzato seleziona e combina le variabili predittive per ottenere pesi di portafoglio ottimali, con l'obiettivo specifico di massimizzare l'utilità, piuttosto che affidarsi a metodi di apprendimento automatico puramente statistici. A livello metodologico, questo approccio integra meccanismi per prevenire il sovradattamento e mantenere i pesi entro limiti predefiniti, senza basarsi sull'ottimizzazione statistica per valutare l'utilità delle variabili predittive nelle decisioni di allocazione.

Applicato ai dati mensili europei, il metodo non parametrico, valutato sulle diverse preferenze dell'investitore attraverso i parametri  $\omega^{max}$ , limite superiore dei pesi prefissato, e  $\gamma$ , coefficiente di avversione al rischio, mostra valori economicamente informativi rispetto ai benchmark. Inoltre i risultati, misurati tramite i criteri di utilità media, Guadagno di rendimento certo equivalente ( $CER_{gain}$ ) e rapporto di Sharpe, sono generalmente superiori rispetto a quelli ottenuti con il metodo basato sulla regressione lineare. In particolare si nota che i criteri di utilità media e  $CER_{gain}$  mostrano generalmente le stesse preferenze nella scelta delle variabili con più potere predittivo e nella valutazioni di quali tra loro non mostrano rilevanza, perchè possiedono valutazioni inferiori

ai criteri di riferimento. Il rapporto di Sharpe in certe situazioni non si allinea con le scelte degli altri due indicatori, specialmente nella valutazione del potere predittivo delle variabili rispetto ai benchmark. Per tale motivo, sembra preferibile basare la scelta delle strategie di investimento sugli indicatori di utilità media e  $CER_{gain}$ , riducendo l'importanza attribuita al rapporto di Sharpe.

Nelle analisi dei risultati finali, l'obiettivo è individuare le strategie che massimizzano i valori dei criteri di valutazione per ciascun profilo di investitore. Dal confronto tra diversi parametri  $\omega^{max}$ , mantenendo costante  $\gamma$ , emergono molte similitudini nella selezione delle variabili con maggiore potere predittivo, e differenze nella valutazione delle capacità predittive dei predittori rispetto ai criteri di riferimento.

Invece, il confronto tra diversi valori di  $\gamma$  con  $\omega^{max}$  costante indica che un aumento della propensione al rischio ( $\gamma = 3$ ) tende a produrre risultati più elevati.

Il confronto tra i metodi viene effettuato in modo puntuale e, complessivamente, nelle analisi sia univariate che multivariate l'approccio non parametrico rispetto a quello basato sulla regressione lineare fornisce risultati migliori, specialmente quando vengono analizzate le variabili economiche ed aziendali. Per tali variabili entrambi i metodi riscontrano performance predittive più elevate rispetto all'intero insieme di predittori analizzati. Per gli indicatori tecnici, invece, non si nota una marcata prevalenza di un metodo sull'altro, però entrambi ritengono che tali variabili non offrono rilevante potere predittivo rispetto ai benchmark, specialmente nell'analisi multivariata.

L'analisi proposta in questa tesi è semplice e di base, aprendo numerose possibilità di sviluppo futuro. Si potrebbe, ad esempio, esplorare l'impiego di funzioni di utilità differenti o adottare uno schema d'investimento più articolato, che includa decisioni tra molteplici attività rischiose, anziché limitarsi a una singola opzione come qui illustrato. Un'ulteriore possibile estensione potrebbe consistere nell'ottimizzare il ribilanciamento del portafoglio, integrando l'effetto dei costi di transazione nel processo di massimizzazione dell'utilità e introducendo così una componente di disturbo nella determinazione ottimale dei pesi di portafoglio.

I confronti tra i due metodi analizzati sono stati effettuati in modo puntuale,

cercando di determinare, utilizzando le stesse variabili, quale approccio producesse il valore migliore. Un altro possibile sviluppo futuro potrebbe consistere nell'utilizzo di metodi statistici per il confronto tra gli indicatori in modo tale da valutare se le differenze osservate tra i vari approcci siano statisticamente significative.



# Appendice A

## Alberi di regressione

Una scelta comune per il *base learner*  $h_m(\mathbf{x}_{t-1})$  nell'algoritmo di gradient boosting è rappresentata dagli alberi di regressione. Questi modelli presentano diversi vantaggi: sono invarianti rispetto a trasformazioni monotone delle variabili esplicative, gestiscono facilmente variabili di diversa natura e la loro rappresentazione matematica è semplice e compatta, grazie all'uso di funzioni a tratti. Inoltre, gli alberi di regressione sono computazionalmente rapidi e, essendo modelli non parametrici, sono in grado di catturare relazioni tra le variabili che semplici modelli lineari non sono in grado di cogliere.

Siano  $(\mathbf{x}_{t-1,1}, \dots, \mathbf{x}_{t-1,p}) \in \mathbb{R}^p$  le variabili predittive e  $\mathbf{y}_t$  la variabile risposta.

Gli alberi di regressioni possono essere rappresentati come una funzione a tratti del tipo:

$$h(\mathbf{x}_{t-1}) = \sum_{j=1}^J c_j I\{\mathbf{x}_{t-1} \in S_j\} \quad (\text{A.1})$$

dove  $S_j$ , con  $j = 1, \dots, J$ , rappresentano le regioni in cui è suddiviso lo spazio delle variabili predittive,  $I\{\cdot\}$  è una variabile indicatrice e  $c_j$  è la costante usata per modellare la variabile dipendente all'interno di ciascuna regione.

Adattare un albero di regressione richiede di effettuare alcune decisioni: determinare il numero di suddivisioni nello spazio  $\mathbb{R}^p$ , definire la posizione dei punti di suddivisione e scegliere il valore costante  $c_j$  da utilizzare per modellare la variabile dipendente in ciascuna regione. Per quanto riguarda quest'ultimo

aspetto, si ricorre alla media dei valori  $\mathbf{y}_t$  che appartengono alla regione  $S_j$ :  
 $\hat{c}_j = \text{mean}(y_t | \mathbf{x}_{t-1} \in S_j)$

Determinare la posizione dei punti di suddivisione è un processo piuttosto complesso che richiede un algoritmo sequenziale. Viene scelta una variabile  $\mathbf{x}_i$ , con  $i = 1, \dots, p$ , e un punto di divisione  $s$ , suddividendo così lo spazio  $\mathbb{R}^p$  in due regioni tramite un iperpiano:

$$S_1(p,s) = \{X | X_p \leq s\} \quad e \quad S_2(p,s) = \{X | X_p > s\} \quad (\text{A.2})$$

Per ciascuna regione, viene calcolato il valore medio  $\hat{c}_j = \text{mean}(y_t | \mathbf{x}_{t-1} \in S_j(p,s))$ , con  $j = 1,2$  e il corrispondente valore di devianza residua:

$$D_j(p,s) = \sum_{y_t, \mathbf{x}_{t-1} \in S_1(p,s)} (y_t - c_1)^2 + \sum_{y_t, \mathbf{x}_{t-1} \in S_2(p,s)} (y_t - c_2)^2 \quad (\text{A.3})$$

ogni possibile coppia di valori  $(p,s)$  viene analizzata e la scelta viene fatta ricadere sulla coppia che massimizza la *caduta in devianza* ovvero la differenza di devianza tra il passo precedente e quello attuale

$$g_j = D_j - D_j(p,s) \quad (\text{A.4})$$

La procedura continua iterativamente suddividendo, in modo ottimale, le regioni costruite al passo precedente, fino a raggiungere una soglia prestabilita di  $g_j$ . L'ottimizzazione in questo modello, da un passo al successivo, è definita *ottimizzazione miofica*; sebbene essa non garantisca la minimizzazione globale della devianza, offre soluzioni ottimali mantenendo un livello di complessità contenuto.

Questa procedura di selezione ottimale ad ogni passo effettua una vera e propria selezione delle variabili più rilevanti per il modello, rendendo l'albero di regressione una scelta adeguata quando il numero di variabili esplicative supera quello delle osservazioni.

La complessità di questo modello è controllata attraverso il numero  $J$  di nodi terminali. Usualmente, dopo la fase di crescita descritta in precedenza, si procede con una fase di potatura in cui viene introdotta una funzione di

penalizzazione da minimizzare rispetto ai nodi  $J$ .

Tuttavia, il nostro algoritmo di gradient boosting presenta una caratteristica distintiva: per ogni iterazione richiede l'utilizzo di un modello di base semplice  $h_m(\mathbf{x})$  che abbia bassa varianza e, di conseguenza, alta distorsione. Distorsione che verrà abbattuta tramite il processo iterativo di aggiornamento delle stime.

Si utilizza quindi un albero semplice, noto come *stump*, che è costituito da soli due nodi terminali ( $J=2$ ), dove ogni nodo contiene almeno 10 osservazioni e che non richiede che venga effettuata alcuna potatura.



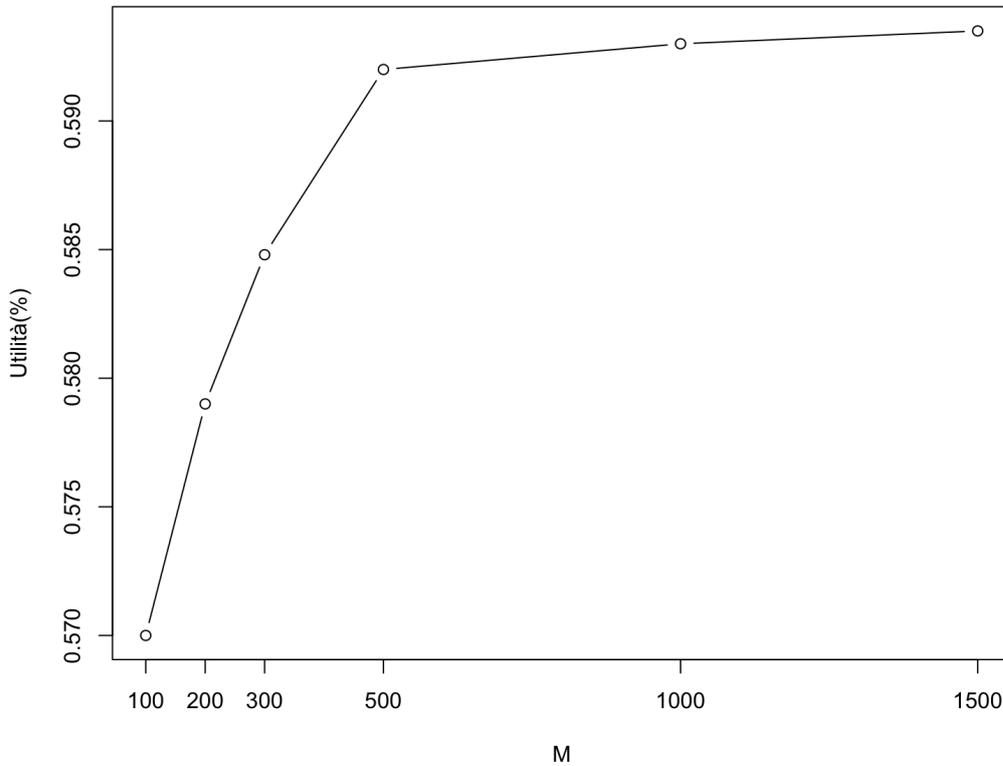
# Appendice B

## Scelta del numero di iterazioni M

Per determinare il numero ottimale di iterazioni M nell'algoritmo di discesa del gradiente, è stata adottata una procedura out-of-sample. Si utilizza una finestra mobile contenente 200 osservazioni, su cui viene allenato il modello di base al passo  $2b$  dell'algoritmo, considerando l'intero insieme di variabili esplicative  $\mathbf{x}_{t-1}$  e i valori  $\gamma = 5$  e  $\omega^{max} = 1$ .

Successivamente, si effettua una previsione utilizzando l'osservazione al passo successivo, e il valore previsto viene impiegato per aggiornare  $F_m(\mathbf{x}_{t-1})$ . Questa procedura viene eseguita per un numero di iterazioni M, e con i risultati  $F_M(\mathbf{x}_{t-1})$  ottenuti, vengono calcolati i pesi  $\omega_t$ . Il processo viene ripetuto spostando la finestra di dati un passo in avanti fino al termine delle osservazioni.

Il procedimento viene eseguito per valori di M pari a 100, 200, 300, 500, 1000 e 1500, e le performance dei pesi ottenuti vengono valutate utilizzando il criterio di *utilità media*. I risultati sono riportati nel grafico seguente:



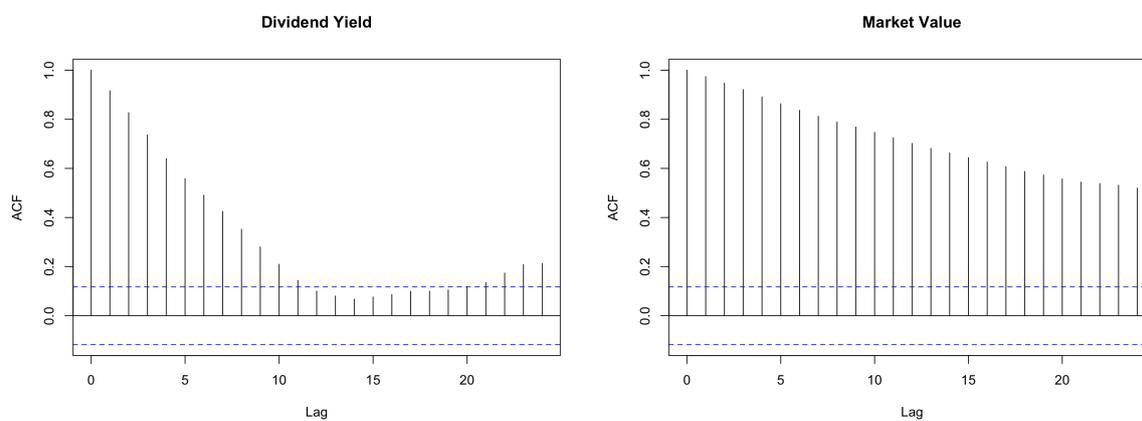
Si può notare come la curva tende a stabilizzarsi dopo  $M=500$ , indicando che un numero maggiore di iterazioni non porta a sostanziali miglioramenti in termini di utilità ma anzi comporta solo un aumento del costo computazionale.

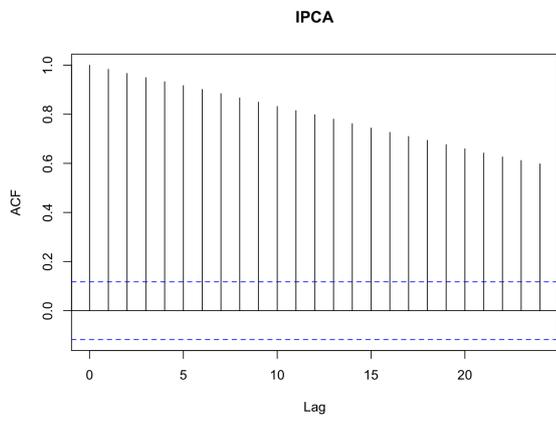
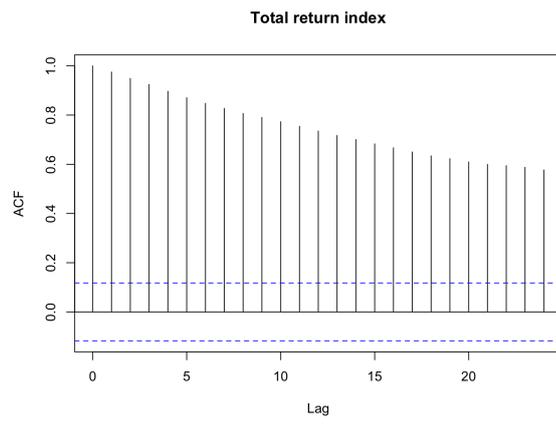
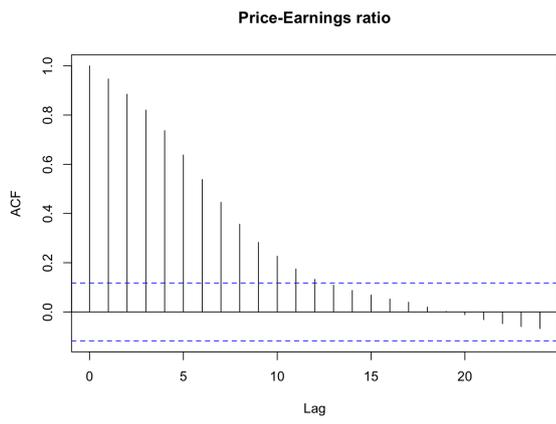
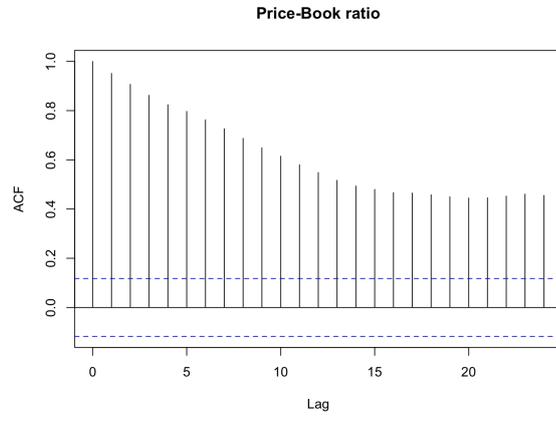
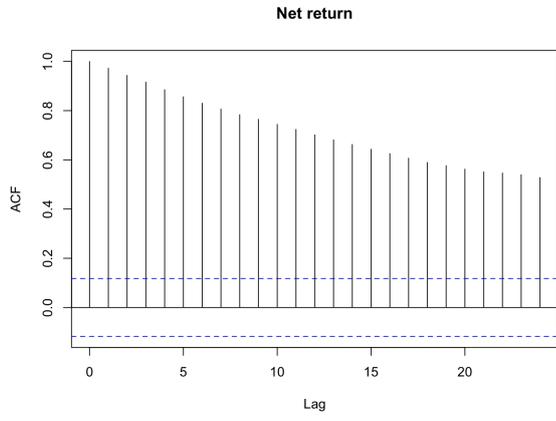
È stata scelta una procedura out-of-sample per la selezione di  $M$  in considerazione della natura dei nostri dati. Trattandosi di serie storiche, l'utilizzo di tecniche di stima-verifica o di convalida incrociata non sarebbe stato appropriato, poiché entrambe richiedono una suddivisione casuale dei dati, la quale comporterebbe una perdita delle informazioni legate alla dipendenza temporale. Inoltre, i fold generati da tali suddivisioni dovrebbero essere indipendenti tra loro, una condizione che non si verifica nel caso delle serie storiche.

# Appendice C

## Analisi Indicatori aziendali ed economici

Si riportano i grafici di autocorrelazione per ciascuna variabile descritta nel *Capitolo 2.1*





Tutte le variabili analizzate mostrano persistenza, ovvero, sono caratterizzate da un'elevata autocorrelazione e memoria lunga. Possono essere adeguatamente descritte tramite processi ARIMA(p,d,q). In particolare le variabili *Dividend Yield*, *Market Value*, *Net Return*, *Price-Earnings Ratio* e *Total Return index* possono essere modellati con un modello Random walk, corrispondente a un modello ARIMA(0,1,0). La variabile *Price-Book ratio* può essere rappresentata da un Random Walk con una componente stagionale autoregressiva, ossia un modello ARIMA(0,1,0)(2,0,0). L' *Indice dei Prezzi al Consumo* invece, è descritto da un modello Arima(1,1,1)(0,1,1), che incorpora una componente autoregressiva, una di media mobile, e una componente stagionale differenziata una volta, con un effetto di media mobile stagionale.

La scelta dei modelli è stata effettuata in base al criterio informativo di Schwartz (BIC)

L'elevata persistenza delle variabili esplicative può distorcere le stime dei coefficienti quando i rendimenti vengono regrediti su di esse secondo il modello  $r_{e,t} = \beta_{i,0} + x_{i,t-1}\beta_{i,1} + \varepsilon_{i,t}$ , per  $i = 1, \dots, p$ .

Per verificare questo effetto, viene analizzata la correlazione tra i residui  $\varepsilon_{i,t}$  e le variabili esplicative  $x_{i,t}$  utilizzando l'intero set di dati:

<b>Variabile</b>	<b>Correlazione</b>
Dividend Yield	-0.092
Market Value	0.158
Net return	0.137
Price-Book ratio	0.191
Price-Earnings ratio	0.123
Total return index	0.116
IPCA	-0.001

Tabella C.1: Correlazioni tra variabili e innovazioni  $\varepsilon_{i,t}$

I valori di correlazione ottenuti non sono elevati e non suggeriscono, quindi, la presenza di una distorsione significativa nei coefficienti.

Data la natura delle variabili presentate, è stato effettuato un test per verificare la presenza di radice unitaria. In particolare, è stato condotto il test di Dickey-Fuller aumentato (ADF) utilizzando un test che include sia un

termine di costante (drift) sia un termine di trend lineare. Di seguito sono riportati i p-value dei risultati:

<b>Variabile</b>	<b>p-value</b>
Dividend Yield	0.014
Market Value	0.329
Net Return	0.585
Price-Book Ratio	0.417
Price-Earnings Ratio	0.010
Total Return Index	0.710
IPCA	0.913

Tabella C.2: Risultati del test di Dickey-Fuller aumentato (ADF)

Analizzando i risultati con un livello di significatività del 5%, si osserva che la maggior parte delle variabili non rifiuta l'ipotesi nulla di presenza di radice unitaria, ad eccezione del *Dividend Yield* e del *Price-Earnings Ratio*.

Per approfondire la verifica della radice unitaria, sono stati effettuati test con differenti configurazioni: introducendo separatamente il trend e il drift, solo il drift o nessuno dei due. I risultati confermano la presenza di radice unitaria per cinque delle variabili analizzate, coerentemente con i risultati precedenti.

Per il *Dividend Yield* e il *Price-Earnings Ratio*, i risultati portano a conclusioni diverse: l'ipotesi nulla di presenza di radice unitaria non viene rifiutata quando si applica il test senza drift né trend, come illustrato nella tabella sottostante, poiché le statistiche test risultano superiori ai valori critici. Questo potrebbe indicare che la serie appaia stazionaria attorno a una struttura con trend o drift, mascherando così la presenza di radice unitaria.

<b>Variabile</b>	<b>Statistica test</b>	<b>Valore critico al 5%</b>
Dividend Yield	-0.5444	-1.95
Price-Earnings Ratio	-1.2848	-1.95

Tabella C.3: Risultati del test di Dickey-Fuller aumentato (ADF) senza drift e trend

# Appendice D

## Grafici

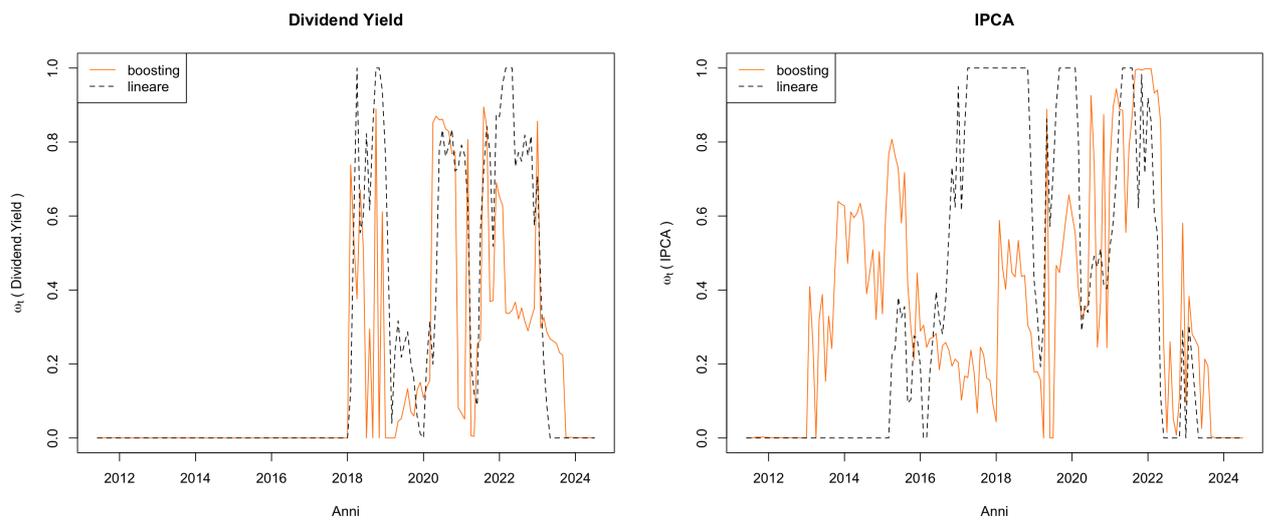


Figura D.1:  $\gamma = 5$  e  $\omega^{max} = 1$

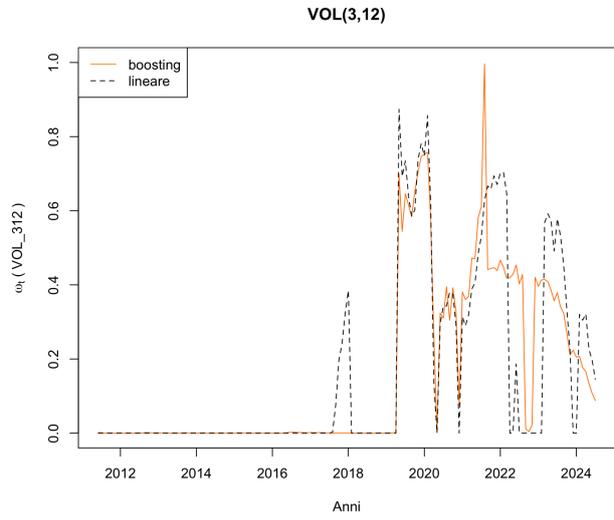
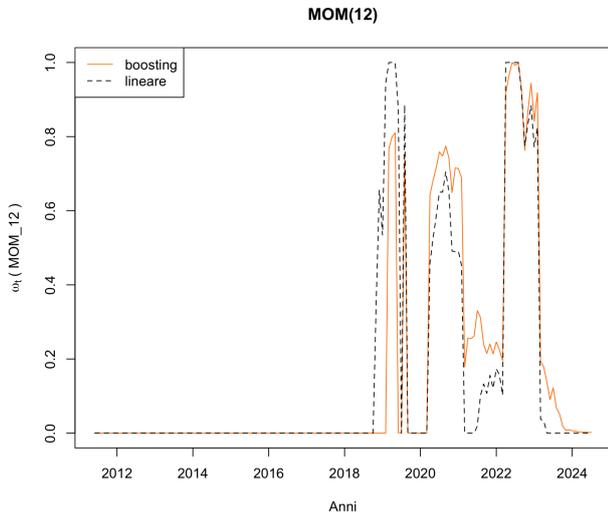


Figura D.2:  $\gamma = 5$  e  $\omega^{max} = 1$

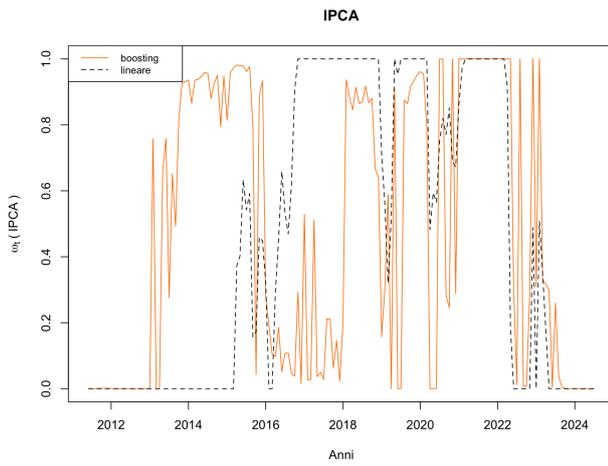


Figura D.3:  $\gamma = 3$  e  $\omega^{max} = 1$

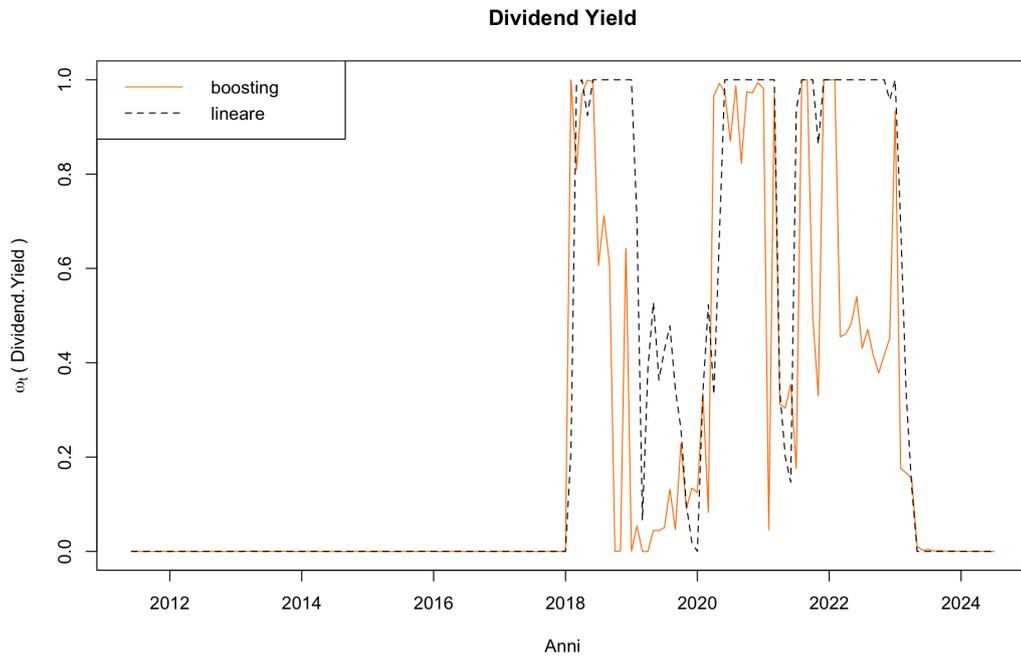


Figura D.4:  $\gamma = 3$  e  $\omega^{max} = 1$

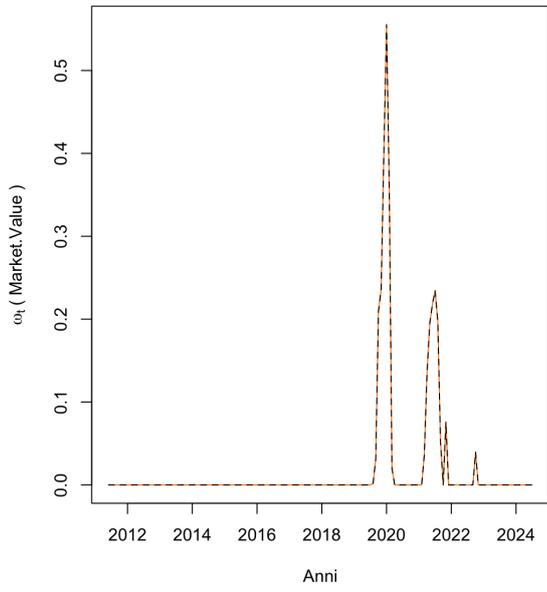


## Appendice E

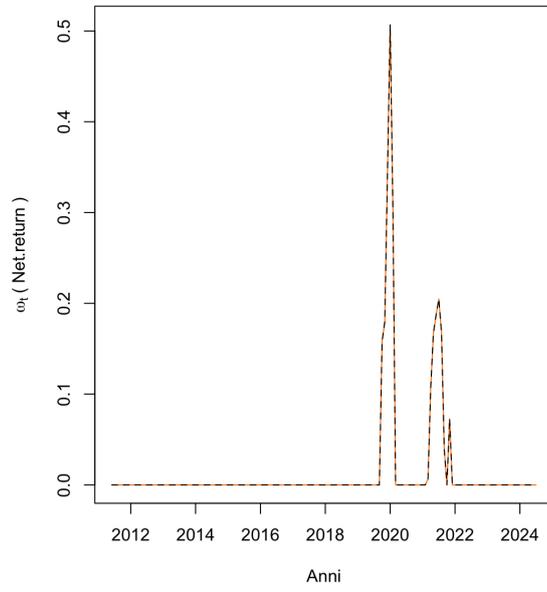
### Confronto grafici per $\omega^{max} = 1$ e $\omega^{max} = 1.5$

Si riportano a confronto i grafici relativi al caso  $\gamma = 5$  delle variabili, che mostrano gli stessi valori dei criteri di valutazione, nelle tabelle (3.2) (3.7), per entrambi i limiti superiori  $\omega^{max} = 1$  e  $\omega^{max} = 1.5$ . La linea arancione rappresenta i pesi troncati a  $\omega^{max} = 1.5$ , mentre la linea tratteggiata nera si riferisce ai pesi troncati a  $\omega^{max} = 1$ . Si osserva come i grafici si uniformino completamente e non raggiungano mai valori superiori a 1.

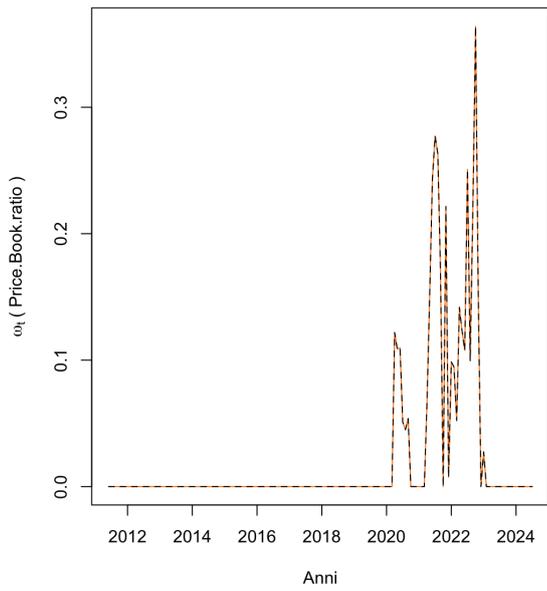
**Market Value**



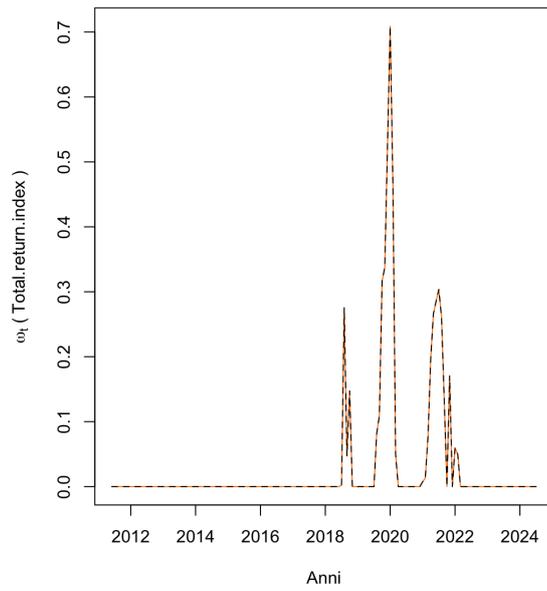
**Net return**



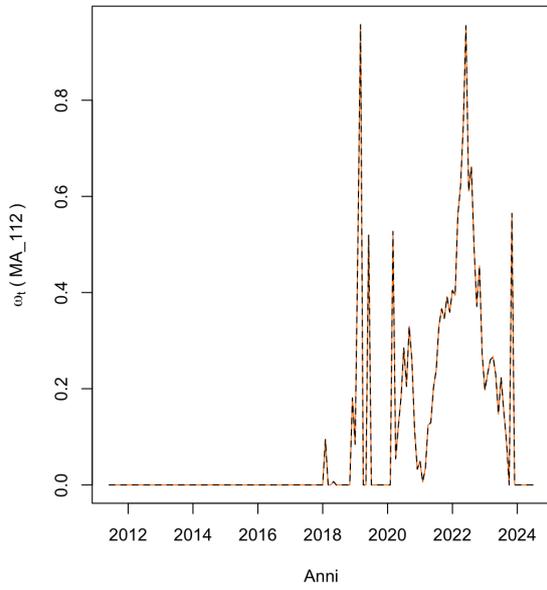
**Price-Book ratio**



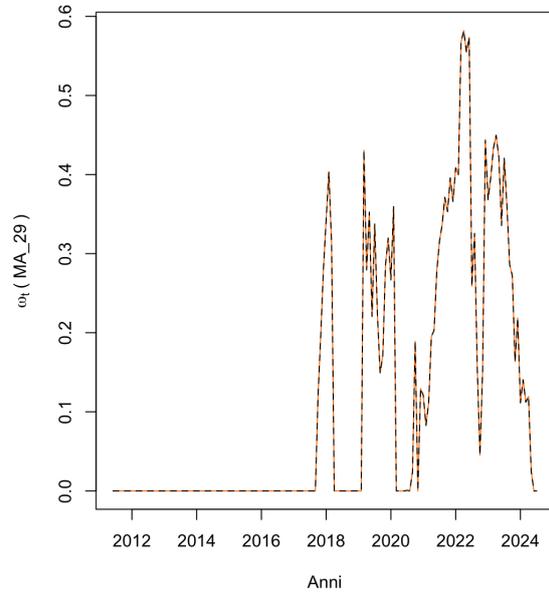
**Total return index**



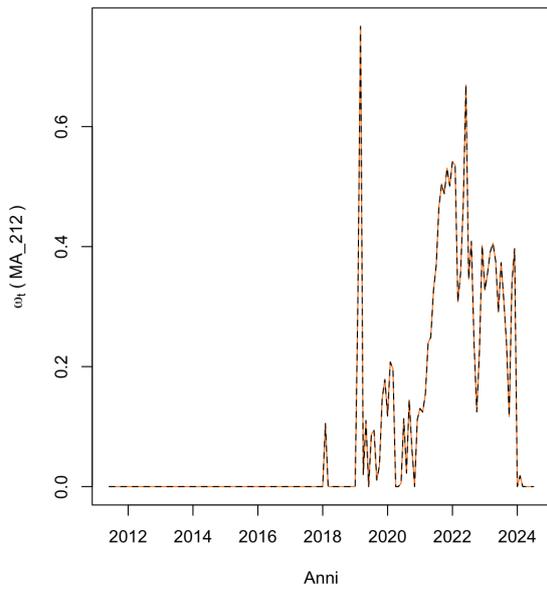
**MA(1,12)**



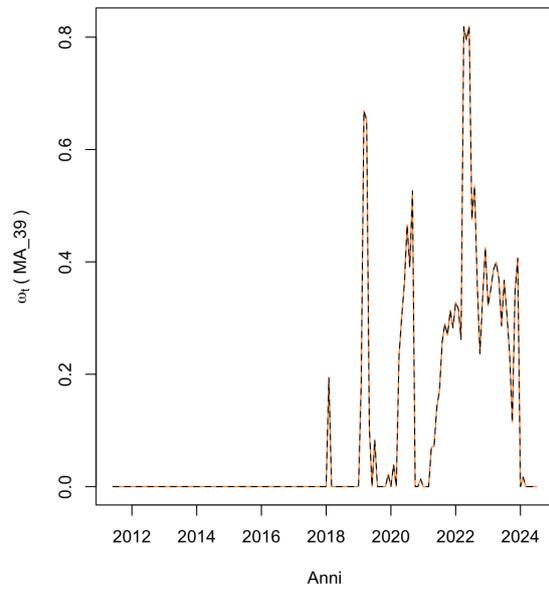
**MA(2,9)**



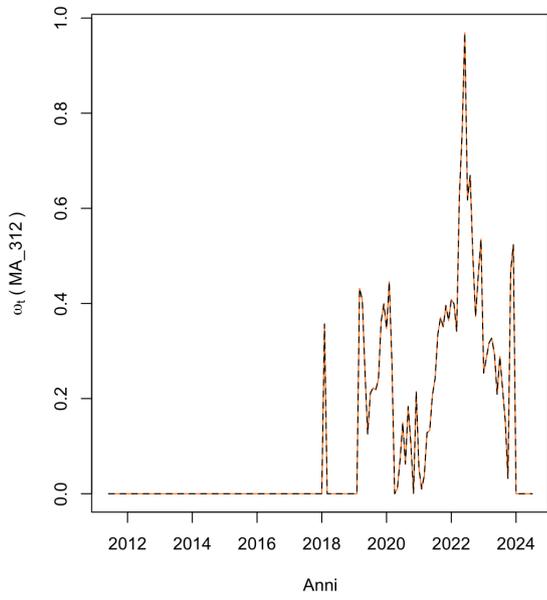
**MA(2,12)**



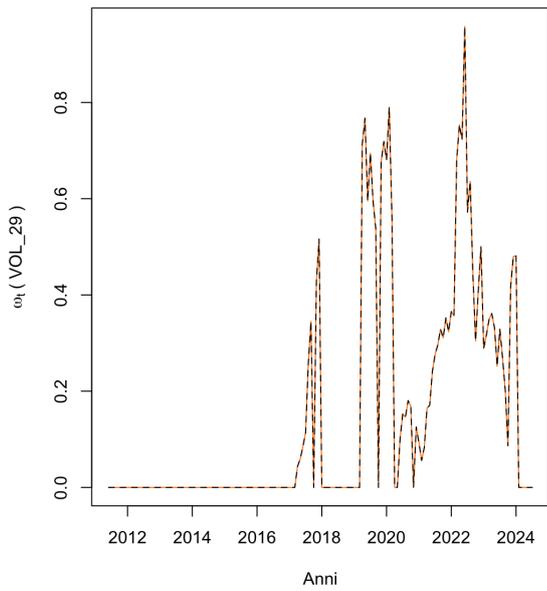
**MA(3,9)**



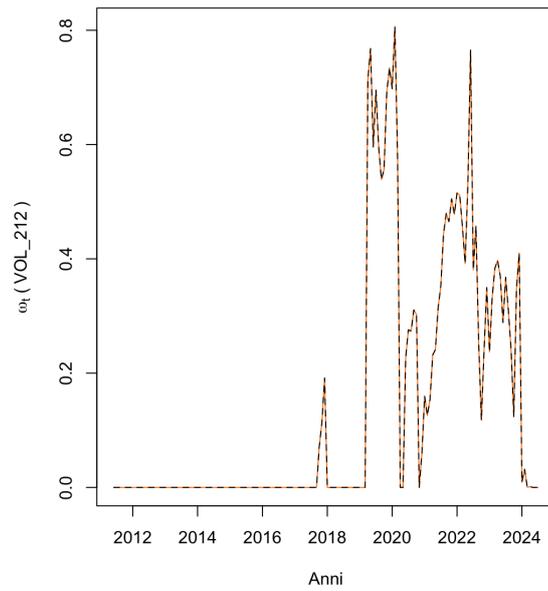
**MA(3,12)**

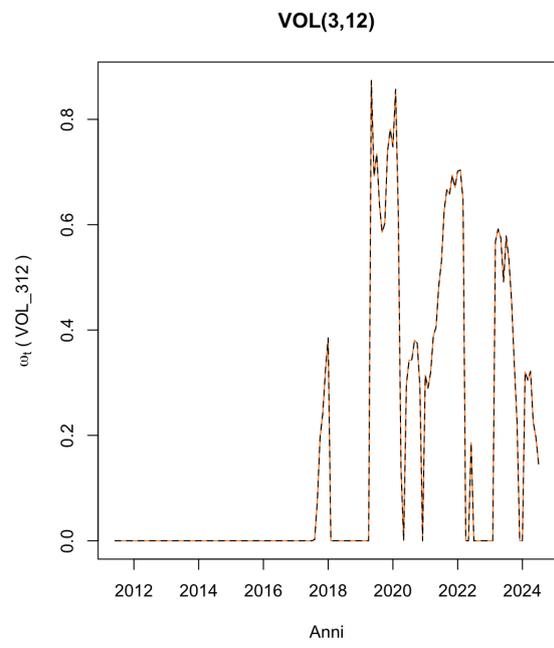
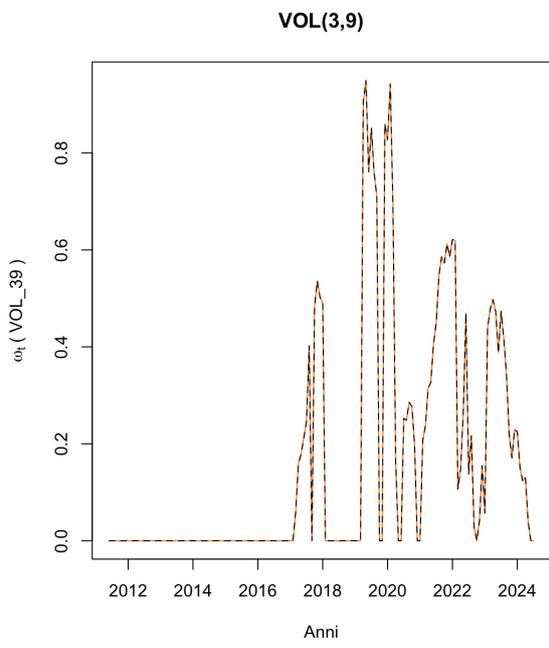


**VOL(2,9)**



**VOL(2,12)**







# Bibliografia

- [1] Yacine Ait-Sahalia and Michael W Brandt. Variable Selection for Portfolio Choice. *Journal of Finance*, 56(4):1297–1351, 2001.
- [2] Francesco Audrino and Fabio Trojani. Accurate Short-Term Yield Curve Forecasting using Functional Gradient Descent. *Journal of Financial Econometrics*, 5:591–623, 2007.
- [3] Adelchi Azzalini, Bruno Scarpa, and Gabriel Walton. *Data analysis and data mining: an Introduction*. Oxford University Press, Oxford ; New York, 2012.
- [4] Gah-Yi Ban, Nouredine El Karoui, and Andrew E. B. Lim. Machine Learning and Portfolio Optimization. *Management Science*, 64(3):1136–1154, 2018.
- [5] Michael W. Brandt. Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach. *The Journal of Finance*, 54(5):1609–1645, October 1999.
- [6] Michael W Brandt and Pedro Santa-Clara. Dynamic Portfolio Selection by Augmenting the Asset Space. *Journal of Finance*, 61:2187–2217, 2006.
- [7] Michael W. Brandt, Pedro Santa-Clara, and Rossen Valkanov. Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns. *Review of Financial Studies*, 22(9):3411–3447, September 2009.
- [8] Peter Bühlmann and Torsten Hothorn. Boosting Algorithms: Regularization, Prediction and Model Fitting. *Statistical Science*, 22:477–505, November 2007.

- [9] John Y Campbell. Predicting Excess Stock Returns out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, 21:1509–1531, 2008.
- [10] Pedro Carmona, Francisco Climent, and Alexandre Momparler. Predicting failure in the U.S. banking sector: An extreme gradient boosting approach. *International Review of Economics & Finance*, 61:304–323, May 2019.
- [11] Tolga Cenesizoglu and Allan Timmermann. Do return prediction models add economic value? *Journal of Banking & Finance*, 36(11):2974–2987, November 2012.
- [12] Peter F Christoffersen. Financial Asset Returns, Direction-of-Change Forecasting, and Volatility Dynamics. *Management Science*, 52:1273–1287, 2006.
- [13] Matei Demetrescu, Iliyan Georgiev, Paulo M.M. Rodrigues, and A.M. Robert Taylor. Testing for episodic predictability in stock returns. *Journal of Econometrics*, 227(1):85–113, March 2022.
- [14] Leland E. Farmer, Lawrence Schmidt, and Allan Timmermann. Pockets of Predictability. *The Journal of Finance*, 78(3):1279–1341, June 2023.
- [15] Jeff Fleming, Chris Kirby, and Barbara Ostdiek. The Economic Value of Volatility Timing. *Journal of Finance*, 56:329–352, 2001.
- [16] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Additive logistic regression: A statistical view of boosting. *Annals of Statistics*, 28(2):337–407, 2000.
- [17] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232, October 2001.
- [18] Jerome H. Friedman. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367–378, February 2002.
- [19] Amit Goyal and Ivo Welch. Predicting the Equity Premium with Dividend Ratios. *Management Science*, 49(5):639–654, 2003.

- [20] Massimo Guidolin. Markov Switching in Portfolio Choice and Asset Pricing Models: A Survey. In David M. Drukker, editor, *Advances in Econometrics*, volume 27, pages 87–178. Emerald Group Publishing Limited, January 2011.
- [21] Massimo Guidolin and Allan Timmermann. Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31(11):3503–3544, November 2007.
- [22] Reuters J.P. Morgan. *RiskMetrics Technical Document - Fourth Edition*. 4 edition, December 1996.
- [23] Raymond Kan and Guofu Zhou. Optimal Portfolio Choice with Parameter Uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3):621–656, September 2007.
- [24] Shmuel Kandel and Robert F Stambaugh. On the Predictability of Stock Returns: An Asset-Allocation Perspective. *Journal of Finance*, 51:385–424, 1996.
- [25] Gordon Leitch and J Ernest Tanner. Economic Forecast Evaluation: Profits Versus the Conventional Error Measures. *American Economic Review*, 81:580–590, 1991.
- [26] Martin Lettau and Stijn Van Nieuwerburgh. Reconciling the Return Predictability Evidence. *Review of Financial Studies*, 21(4):1607–1652, July 2008.
- [27] Wei-Yin Loh and Yu-Shan Shih. Split selection methods for classification trees. *Statistica Sinica*, 7(4):815–840, 1997.
- [28] Verbeek Marno Marquering Wessel. The Economic Value of Predicting Stock Index Returns and Volatility. *The Journal of Financial and Quantitative Analysis*, 39(2):407–429, 2004.
- [29] Antonio Mele. Asymmetric stock market volatility and the cyclical behavior of expected returns. *Journal of Financial Economics*, 86(2):446–478, November 2007.

- [30] Stefan Mittnik, Nikolay Robinzonov, and Martin Spindler. Stock market volatility: Identifying major drivers and the nature of their impact. *Journal of Banking & Finance*, 58:1–14, September 2015.
- [31] Christopher J. Neely, David E. Rapach, Jun Tu, and Guofu Zhou. Forecasting the Equity Risk Premium: The Role of Technical Indicators. *Management Science*, 60(7):1772–1791, July 2014.
- [32] Lauri Nevasalmi and Henri Nyberg. Moving Forward from Predictive Regressions: Boosting Asset Allocation Decisions. *SSRN Electronic Journal*, January 2021.
- [33] M Hashem Pesaran and Allan Timmermann. Predictability of Stock Returns: Robustness and Economic Significance. *Journal of Finance*, 50:1201–1228, 1995.
- [34] Davide Pettenuzzo, Allan Timmermann, and Rossen Valkanov. Forecasting stock returns under economic constraints. *Journal of Financial Economics*, 114(3):517–553, December 2014.
- [35] David Rapach and Guofu Zhou. Forecasting Stock Returns. In *Handbook of Economic Forecasting*, volume 2, pages 328–383. Elsevier, 2013.
- [36] David E. Rapach, Jack K. Strauss, and Guofu Zhou. Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. *Review of Financial Studies*, 23(2):821–862, February 2010.
- [37] Alberto Rossi. Predicting stock market returns with machine learning. August 2018.
- [38] Alberto G. Rossi and Allan Timmermann. Modeling Covariance Risk in Merton’s ICAPM. *The Review of Financial Studies*, 28(5):1428–1461, May 2015.
- [39] E. Sentana. Least Squares Predictions and Mean-Variance Analysis. *Journal of Financial Econometrics*, 3(1):56–78, January 2005.
- [40] Robert F Stambaugh. Predictive regression. *Journal of Financial Economics*, 54:375–421, 1999.

- [41] Ivo Welch and Amit Goyal. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, 21(4):1455–1508, July 2008.
- [42] Yexiao Xu. Small levels of predictability and large economic gains. *Journal of Empirical Finance*, 11(2):247–275, March 2004.
- [43] Yaojie Zhang, Yu Wei, Feng Ma, and Yongsheng Yi. Economic constraints and stock return predictability: A new approach. *International Review of Financial Analysis*, 63:1–37, May 2019.