

Bachelorarbeit  
im Bachelorstudiengang Wirtschaftspsychologie (B.Sc.)  
an der Hochschule für angewandte Wissenschaften Neu-Ulm

**Thema:**  
**Ökonometrische Prognosemodelle für Anleiherenditen: Ein empirischer Vergleich für 10-jährige US-Staatsanleihen**

Erstkorrektor: Hr. Prof. Dr. Thorsten Neumann  
Zweitkorrektor: Hr. Prof. Dr. Markus Wilhelm

Verfasser: Moritz Schwally (Matrikel-Nr.: 294724)

Thema erhalten: 26.06.2025  
Arbeit abgeliefert: 27.10.2025

## Abstract

Diese Arbeit untersucht, welche ökonometrischen Prognosemodelle die höchste Vorhersagegenauigkeit für die Rendite 10-jähriger US-Staatsanleihen erzielen und inwieweit sich diese Prognosen in profitable Handelsstrategien umsetzen lassen. Ausgangspunkt ist die zentrale Rolle der 10-jährigen Treasury-Rendite als globaler Referenzzins sowie das Spannungsfeld zwischen Markteffizienz und fundamentaler Analyse. Anhand monatlicher Daten von 1977 bis 2025 werden 27 Modelle geschätzt, darunter univariate ARIMA-, multivariate Regressions-, Fehlerkorrektur- (ECM) und hybride ARIMAX-Modelle. Die Prognosegüte wird mittels MAE, Hit Ratio und Weighted Hit Ratio bewertet, die ökonomische Verwertbarkeit durch simulierte Long/Short-Strategien auf den 10-Year-Treasury-Future überprüft. Die Ergebnisse zeigen, dass insbesondere das ECM\_2-Modell (Sharpe = 0,683; Rendite = 3,91 %) die beste risikoadjustierte Performance erzielt, während Buy-and-Hold nur 0,12 % Rendite bei 24,48 % Drawdown erreicht. Die Einbeziehung fundamentaler Variablen verbessert somit Prognosegüte und Effizienz signifikant. Die Ergebnisse weisen auf temporäre Ineffizienzen in den Zinsmärkten hin, die durch fundamentale Modelle systematisch genutzt werden können und damit Grenzen der schwachen Form der Effizienzmarkthypothese aufzeigen. Die Arbeit leistet einen Beitrag, indem sie die Prognosegüte verschiedener Modellklassen empirisch vergleicht, deren ökonomische Verwertbarkeit in Handelsstrategien überprüft und damit zentrale Forschungsfragen zur Vorhersagbarkeit und praktischen Relevanz langfristiger Zinssätze beantwortet.

Key words: Zinsprognose, ARIMA, Error-Correction-Model, ARIMAX, US-Treasury-Rendite, Handelsstrategie

## Inhaltsverzeichnis

Abstract.....	II
Abbildungsverzeichnis .....	V
Tabellenverzeichnis .....	VI
Abkürzungsverzeichnis .....	VII
1 Einleitung.....	9
1.1 Entstehung des US-Treasury-Marktes und –Instituts .....	10
1.2 Relevanz und Problemstellung .....	11
2 Methodische und theoretische Grundlagen der Zinsprognose .....	15
2.1 Grundlagen der Zeitreihenanalyse .....	18
2.2 Univariate Prognosemodelle: Die ARIMA-Klasse .....	20
2.3 Multivariate Prognosemodelle: Lineare Regression, ARIMAX und Fehlerkorrekturmodelle (ECM) .....	22
2.4 Gütekriterien zur Evaluierung von Prognosen .....	25
2.5 Anwendung in Handelsstrategien .....	31
2.6 Zielsetzung und Forschungsfragen .....	31
2.7 Stand der Forschung .....	32
3 Datenbasis und empirisches Vorgehen .....	36
3.1 Datenakquise und -beschreibung .....	36
3.2 Datenaufbereitung und Stationaritätstests .....	40
3.3 Modellspezifikation und Prognosestrategie .....	42
4 Empirische Ergebnisse .....	50
4.1 Univariate ARIMA-Modelle: Baseline-Performance .....	50
4.2 Fundamentale Regressionsmodelle: Integration makroökonomischer Informationen .....	52
4.3 Fehlerkorrekturmodelle (ECM): Langfristige Gleichgewichtsbeziehungen .....	54
4.4 Hybride ARIMAX-Modelle: Integration von autoregressiven, fundamentalen und exogenen Dynamiken.....	55
4.5 Gesamtvergleich und Modellselektion .....	57
4.6 Handelsstrategien: Evaluation der Profitabilität .....	59
5 Diskussion und Fazit.....	69
5.1 Zusammenfassung der zentralen Ergebnisse .....	69

---

5.2	Beantwortung der Forschungsfragen .....	70
5.3	Theoretische Einordnung und Implikationen für die Effizienzmarkthypothese.....	71
5.4	Kritische Würdigung und Limitationen .....	72
5.5	Ausblick für Forschung und Praxis .....	73
	Literaturverzeichnis.....	76

---

## Abbildungsverzeichnis

Abbildung 1: Entwicklung der Rendite 10-jähriger US-Treasuries (1977–2025).....	9
Abbildung 2: Typische Formen der US-Zinsstrukturkurve (normal, flach, invers) in verschiedenen Marktphasen.....	12
Abbildung 3: Darstellung der Modell-Performance der zehn besten Prognosemodelle (MAE vs. Weighted Hit Ratio). ....	58
Abbildung 4: Risiko-Rendite-Profil der zehn besten Trading-Strategien (Annualisierte Volatilität vs. Annualisierte Rendite). ....	64
Abbildung 5: Performance-Vergleich der modellbasierten Handelsstrategien im Zeitraum 2006–2025 auf Monatsbasis.....	67

## Tabellenverzeichnis

<b>Tabelle 1:</b> Übersicht der transformierten Variablen und angewandten Transformationen	38
<b>Tabelle 2:</b> Übersicht der univariaten Prognosemodelle (ARIMA-Klasse)	44
<b>Tabelle 3:</b> Übersicht der fundamentalen Regressionsmodelle (multivariat)	45
<b>Tabelle 4:</b> Übersicht der Error-Correction-Modelle (ECM-Klasse)	46
<b>Tabelle 5:</b> Übersicht der hybriden ARIMAX-Modelle (kombinierte Zeitreihen- und Fundamentaldatenansätze)	47
<b>Tabelle 6:</b> Übersicht der verwendeten Modellklassen und Spezifikationen	48
<b>Tabelle 7:</b> Prognosegüte der univariaten Modelle (ARIMA/MA-Klasse)	50
<b>Tabelle 8:</b> Prognosegüte der fundamentalen Regressions- und ECM-Modelle	52
<b>Tabelle 9:</b> Prognosegüte der Fehlerkorrekturmodelle (ECM-Klasse)	54
<b>Tabelle 10:</b> Prognosegüte der ARIMAX-Modelle	56
<b>Tabelle 11:</b> Gesamtvergleich der Prognosegüte aller Modellklassen	57
<b>Tabelle 12:</b> Performancekennzahlen der modellbasierten Handelsstrategien	59
<b>Tabelle 13:</b> Performancekennzahlen der fundamentalen und ECM-basierten Handelsstrategien	60
<b>Tabelle 14:</b> Vergleich der Performancekennzahlen aller Modellklassen	62
<b>Tabelle 15:</b> Optimierte Gewichte des Minimum-Variance-Portfolios	65
<b>Tabelle 16:</b> Vergleich der Portfolio-Performance modellbasierter Strategien	66

## Abkürzungsverzeichnis

ACF	Autokorrelationsfunktion (Autocorrelation Function)
ADF	Augmented Dickey-Fuller (Test)
AIC	Akaike Information Criterion
AR	Autoregressiv (Autoregressive)
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Exogenous Variables
BIC	Bayesian Information Criterion
COR	Korrelation (Correlation)
CPI	Consumer Price Index (Verbraucherpreisindex)
CPIAUCSL	Consumer Price Index for All Urban Consumers
DGS2	2-jährige US-Staatsanleihenrendite
DGS10	10-jährige US-Staatsanleihenrendite
DNS	Dynamic Nelson-Siegel (Modell)
ECM	Error Correction Model (Fehlerkorrekturmodell)
ECT	Error Correction Term (Fehlerkorrekturterm)
EHTS	Expectations Hypothesis of the Term Structure
EMH	Efficient Market Hypothesis (Effizienzmarkthypothese)
FAVAR	Factor-Augmented Vector Autoregression
FEDFUNDS	Federal Funds Rate (Leitzins der Federal Reserve)
FOMC	Federal Open Market Committee
FRED	Federal Reserve Economic Data
HR	Hit Ratio (Trefferquote)
INDPRO	Industrial Production Index (Industrieproduktionsindex)
M2SL	Geldmenge M2
MA	Moving Average (Gleitender Durchschnitt)

---

MAE	Mean Absolute Error (Mittlerer absoluter Fehler)
MDD	Maximum Drawdown (Maximaler Drawdown)
ML	Maximum Likelihood
MSE	Mean Squared Error (Mittlerer quadratischer Fehler)
NBER	National Bureau of Economic Research
OLS	Ordinary Least Squares (Methode der kleinsten Quadrate)
PACF	Partielle Autokorrelationsfunktion
PPI	Producer Price Index (Erzeugerpreisindex)
PPIACO	Producer Price Index for All Commodities
Rol.	Rolling (rollierend)
Rek.	Rekursiv
SP500	S&P 500 Index
UNRATE	Arbeitslosenquote (Unemployment Rate)
US	United States (Vereinigte Staaten)
VAR	Vector Autoregression
WHR	Weighted Hit Ratio (Gewichtete Trefferquote)

# 1 Einleitung

Seit Jahrzehnten stehen Zinsentwicklungen im Zentrum wirtschaftlicher, geldpolitischer und wissenschaftlicher Diskussionen. Kaum eine Kennziffer fasst die Erwartungen über Inflation, Wachstum und Stabilität so prägnant zusammen wie die Rendite langfristiger Staatsanleihen. Insbesondere die 10-jährige US-Treasury-Note gilt als Barometer der globalen Finanzarchitektur, ein Indikator, an dem sich wirtschaftlichen Trends ablesen lassen (Guerard, 2013, ch. 2). Ihre Entwicklung erzählt zugleich die Geschichte der modernen Geldpolitik, geprägt von Phasen hoher Inflation, wirtschaftlicher Stabilisierung und historisch niedriger Zinsen.

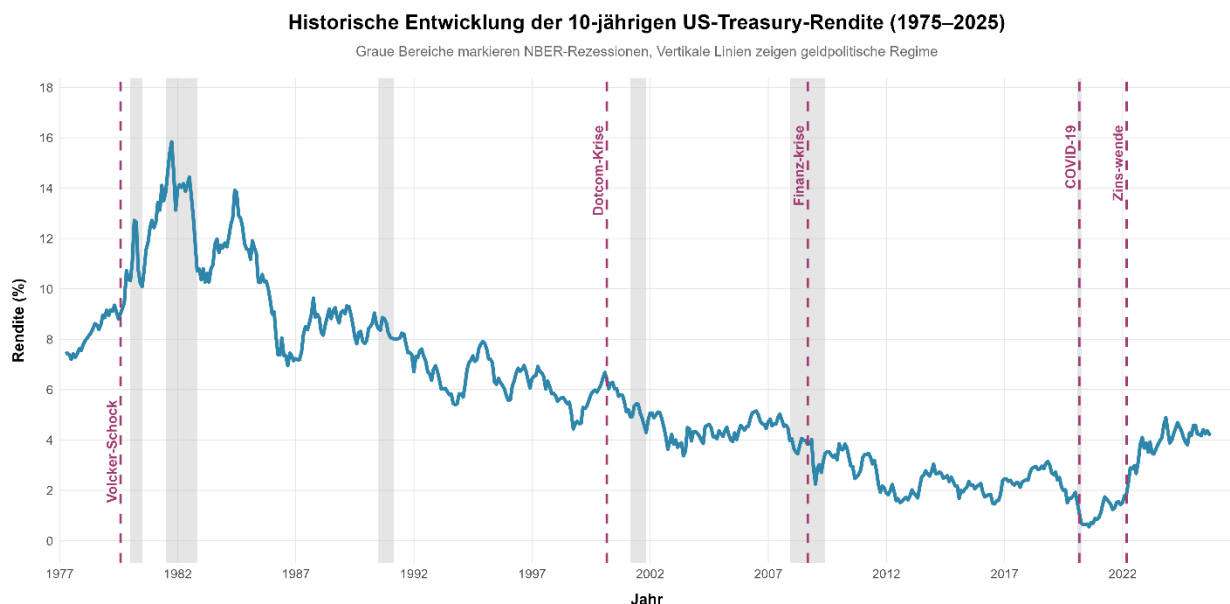


Abbildung 1: Entwicklung der Rendite 10-jähriger US-Treasuries (1977–2025); Quelle: eigene Darstellung auf Basis von FRED-Daten (DGS10), National Bureau of Economic Research (NBER)-Rezessionsphasen grau hinterlegt, Vertikale Linien zeigen geldpolitische Regime.

Ein Blick in die Historie verdeutlicht die Wandlungsfähigkeit dieses Instruments (Abb. 1). Anfang der 1980er-Jahre, im Umfeld des sogenannten Volcker-Schocks, erreichten die Renditen 10-jähriger US-Staatsanleihen zeitweise Werte von über 15 % (Abbildung), was den Einfluss sehr hoher Inflationserwartungen und einer restriktiven Geldpolitik widerspiegelt (Duffee, 2012). In den folgenden Jahrzehnten entwickelten sich die Zinsen in verschiedenen Regimen, wobei sie sich in den 1990er- und frühen 2000er-Jahren auf deutlich moderaterem Niveau bewegten, da die Inflation zurückging und das Wirtschaftswachstum an Stabilität gewann (Ang & Piazzesi, 2003; Palazzo & Nobili, 2010). Nach der globalen Finanzkrise 2008 begann dann eine Phase beispiellos niedriger Zinssätze, teils unter 2 % langfristig, bedingt durch expansive Geldpolitik, Null- bzw. Negativzinsen sowie umfangreiche quantitative Lockerungsmaßnahmen (Müller et al., 2022).

Spätestens seit der COVID-19-Pandemie und der anschließenden Inflationswelle 2021–2023 hat die Dynamik in der US-Zinslandschaft eine neue Dimension erreicht. Auf eine Phase sehr niedriger Zinsen folgte der schnellste Zinsstraffungszyklus seit vier Jahrzehnten, wodurch die langfristigen Renditen wieder stark anzogen (Wang et al., 2025). Gleichzeitig spiegelt die Bewegung der Zinsen nicht mehr nur Inflationserwartungen und Wachstumserwartungen wider, sondern zunehmend auch Unsicherheit über institutionelle und strukturelle Rahmenbedingungen (Müller, 2013, ch. 1).

Parallel dazu ist die Rolle der US-Treasuries als globaler Stabilitätsanker stärker denn je in den Fokus gerückt, da der US-Treasuries Markt den größten und liquidesten Staatsanleihemarkt der Welt bildet (Gürkaynak et al., 2006; Chaboud et al., 2025). In Phasen finanzieller Turbulenzen, ob während der Eurokrise, der Pandemie oder angesichts geopolitischer Spannungen, fungieren sie als „Safe Asset“, also als risikoarme Anlageklasse, die Anlegern Kapitalerhalt und Liquidität bietet (Palazzo & Nobili, 2010). Gleichwohl zeigte die Marktstörung im März 2020, dass selbst dieser Markt nicht gegen Liquiditätsengpässe geschützt ist und die Stabilität des globalen Finanzsystems zunehmend von der Interventionsfähigkeit der Federal Reserve abhängt (Sahr, 2024; Müller et al., 2022).

Um die heutige Dominanz der US-Treasuries im globalen Finanzsystem zu verstehen, ist ein Blick auf ihre institutionellen Ursprünge notwendig. Ihre Entwicklung von einem reinen Schuldenverwaltungsinstrument hin zu einem zentralen Pfeiler der Weltfinanzordnung reicht bis in die Anfänge der amerikanischen Republik zurück.

## **1.1 Entstehung des US-Treasury-Marktes und –Instituts**

Die institutionelle Grundlage des heutigen US-Staatsanleihe-marktes lässt sich zurückführen auf die Gründung des US-Finanzministeriums (United States Department of the Treasury) im Jahr 1789, das zum Ziel hatte, die im Unabhängigkeitskrieg entstandenen Schulden zu verwalten und die Zahlungsfähigkeit der jungen Nation zu sichern. Über die Jahrzehnte entwickelte sich die US-Schuldenaufnahme und -verwaltung systematisch weiter, von vereinzelt begebenen Staatsanleihen zur Finanzierung einzelner Ausgaben hin zu einem systematischen Markt, der langfristige Anleihen, Auktionen und ein priorisiertes Schuldenmanagement umfasst (Garbade, 2012, ch. Introduction).

Ein Meilenstein war die Entwicklung des Markts für auflegbare langfristige Staatsanleihen im frühen 20. Jahrhundert. In seinem Werk „Birth of a Market: The U.S. Treasury Securities Market from the Great War to the Great Depression“ beschreibt Garbade (2012), wie sich zwischen 1917 und 1939 schrittweise vier wesentliche Reformen im Schuldenmanagement des US-Treasury etablierten. Dazu zählen die Einführung regelmäßiger und planbarer Auktionen, die enge Verknüpfung von Kassen- und Schuldenverwaltung, die Lockerung gesetzlicher Restriktionen für das Management der Staatsverschuldung sowie die institutionelle Herausbildung eines modernen und effizienten Marktes für Staatsanleihen

(Garbade, 2012, ch. 1). Diese Entwicklungen führten den US-Treasury-Markt aus seiner frühen Nischenrolle zu einem modernen, liquiden und systemrelevanten Markt.

Parallel dazu förderte die technische Weiterentwicklung des Handelssystems die Professionalisierung des Marktes. In den 1960er- und 1970er-Jahren begann die schrittweise Umstellung von papierbasierten Staatsanleihen auf ein Buch-Entry-System („book-entry securities“), was sowohl die Kosten senkte als auch die Handelbarkeit und Marktliquidität deutlich verbesserte (Garbade, 2004). Die Kombination institutioneller Reformen und technischer Innovationen machte den US-Treasury-Markt zum größten und liquidesten Staatsanleihemarkt der Welt, eine zentrale Referenz und Benchmark für globale Kapitalflüsse und Risikobewertungen.

Garbade (2012) zeigt, dass sich die Komplexität und internationale Bedeutung dieses Marktes aus einer Reihe tiefgreifender institutioneller und struktureller Entwicklungen ergaben. Nach der Etablierung regelmäßiger Auktionen und der Standardisierung von Anleihebedingungen in den 1920er-Jahren führte insbesondere die enge Verzahnung von Primär- und Sekundärmarkt, die wachsende Rolle privater Finanzinstitute als Intermediäre sowie der Ausbau einer effizienten Marktinfrastruktur zu einer nachhaltigen Professionalisierung des Handels (Garbade, 2012, ch. 15). Die Öffnung des Marktes für institutionelle Anleger, die Entwicklung eines aktiven Repurchase-(Repo-)Marktes und die zunehmende Einbindung internationaler Kapitalströme im Zuge der Nachkriegszeit bewirkten schließlich, dass sich der ursprünglich nationale Staatsanleihemarkt zu einem hochliquiden, global vernetzten Finanzsystem entwickelte (Garbade, 2012, ch. 13).

Damit wird deutlich, dass sich aus der ursprünglich auf die Finanzierung kriegsbedingter Staatsausgaben im 18. und 19. Jahrhundert ausgerichteten Aufgabe über die Jahrzehnte ein komplexes Markt- und Institutionensystem entwickelte, das heute das Rückgrat der globalen Finanzarchitektur bildet. Das Zusammenspiel aus historischer Entwicklung, institutioneller Gestaltung und Marktdynamik macht den US-Treasury-Markt zu einem zentralen Untersuchungsgegenstand für das Verständnis langfristiger Zinsentwicklungen.

## **1.2 Relevanz und Problemstellung**

Aus der historischen und institutionellen Entwicklung geht hervor, dass der US-Treasury-Markt eine zentrale Rolle im internationalen Finanzsystem spielt und wesentlich die globale Preisbildung und Kapitalverteilung prägt. Vor diesem Hintergrund kommt insbesondere den 10-jährigen US-Staatsanleihen („10-Year Treasury Notes“) eine herausragende Bedeutung zu, da sie als eines der zentralen Referenzinstrumente der globalen Finanzmärkte gelten. Ihre Rendite gilt als der wichtigste risikofreie Zinssatz der Welt und dient als Benchmark für die Preisbildung beinahe aller anderen Vermögenswerte, von Hypothekenzinsen über Unternehmensanleihen bis hin zu komplexen Derivaten (Ang & Piazzesi, 2003). Aufgrund der hohen Kreditwürdigkeit der Vereinigten Staaten werden US-Staatsanleihen als nahezu

risikolos angesehen, wodurch ihre Renditen eine zentrale Rolle bei der Bewertung von Finanzierungskosten und Kapitalströmen einnehmen (Guidolin & Ionta, 2025).

Darüber hinaus fungiert die 10-Jahres-Rendite als makroökonomischer Indikator, der die Erwartungen von Investoren bezüglich Wachstum, Inflation und Geldpolitik reflektiert (Koroleva & Kopeykin, 2022). Steigende Renditen signalisieren häufig Optimismus über zukünftiges Wachstum, während fallende Renditen Unsicherheit und Kapitalflucht in sichere Anlagen anzeigen (Joshi, o.J.). Diese Doppelfunktion als Preisanker und Konjunkturbarometer verleiht der exakten Prognose der 10-jährigen Treasury-Rendite eine erhebliche wissenschaftliche und praktische Relevanz (Diebold, Rudebusch & Aruoba, 2006).

Diese Relevanz zeigt sich besonders in den jüngsten makroökonomischen Entwicklungen. Nach einer Phase historisch niedriger Zinsen infolge der expansiven Geldpolitik der Federal Reserve (2020–2022) stieg die 10-Jahres-Rendite ab 2023 infolge der aggressiven Zinsanhebungen zur Inflationsbekämpfung deutlich an. Diese Entwicklung führte zu einer anhaltend inversen Zinsstrukturkurve, einem klassischen Rezessionssignal und veränderte die Dynamik globaler Kapitalflüsse erheblich (Wang et al., 2025). Damit wurde erneut veranschaulicht, wie sensibel langfristige Zinssätze auf geldpolitische Erwartungen und makroökonomische Rahmenbedingungen reagieren.

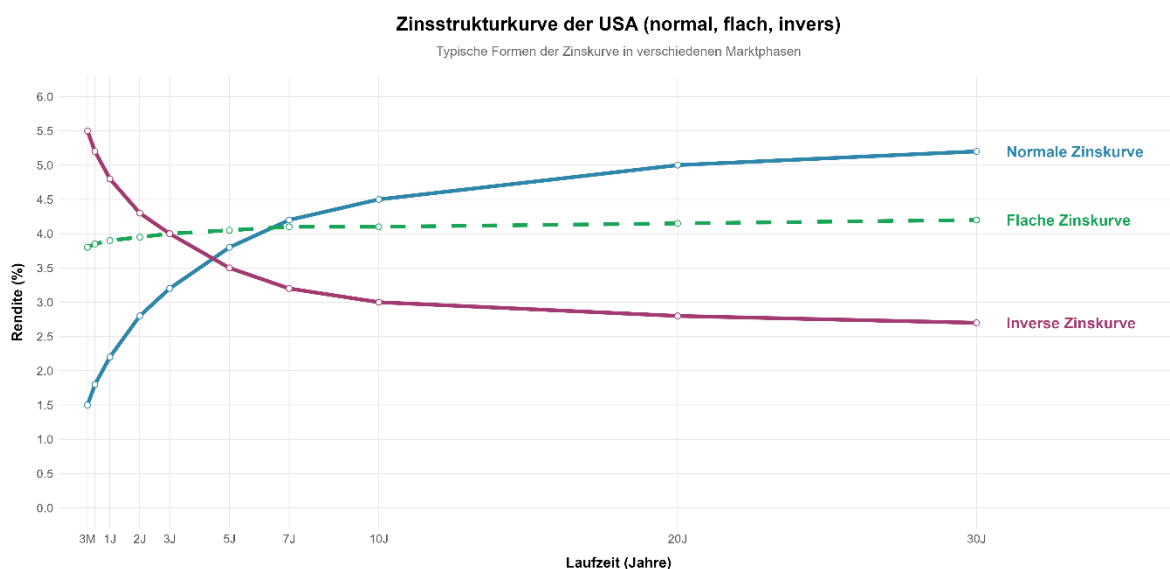


Abbildung 2: Typische Formen der US-Zinsstrukturkurve (normal, flach, invers) in verschiedenen Marktphasen; Quelle: Eigene Darstellung auf Basis theoretischer Zinsverläufe.

Die jüngsten Beobachtungen zeigen, dass Zinssätze nicht isoliert entstehen, sondern das Ergebnis eines komplexen Zusammenspiels zwischen geldpolitischen Entscheidungen, makroökonomischen Fundamentalfaktoren und Markterwartungen sind. Zur theoretischen Erklärung dieses Zusammenhangs liefert die Erwartungstheorie der Zinsstruktur einen zentralen Ansatz. Sie geht davon aus, dass sich langfristige Zinssätze aus den Erwartungen über zukünftige Kurzfristzinsen ergeben, ergänzt um eine Termprämie, die Anleger für das

Risiko längerer Bindung kompensiert (Diebold, Rudebusch & Aruoba, 2006; Guidolin & Ionta, 2025). Damit werden langfristige Renditen zu Verdichtungspunkten kollektiver Markterwartungen, eine Schnittstelle zwischen realwirtschaftlicher Fundamentaldynamik und psychologischer Marktstimmung.

Gleichzeitig offenbart sich in diesem Zusammenhang ein grundlegendes theoretisches Spannungsfeld. Nach der Effizienten-Markt-Hypothese (Fama, 1970) spiegeln Renditen bereits alle verfügbaren Informationen wider, sodass kein Prognosemodell eine systematische Überlegenheit erzielen könnte. Trotzdem zeigt die Erwartungstheorie der Zinsstruktur, dass langfristige Zinssätze aus den Erwartungen über zukünftige Kurzfristzinsen entstehen, eine theoretische Grundlage, die Prognosemodelle überhaupt erst rechtfertigt.

Diese zwei theoretischen Perspektiven führen zu einem zentralen Spannungsfeld, wenn Märkte effizient sind und alle verfügbaren Informationen sofort in Preisen reflektiert werden (Fama, 1970), dürfte es keine systematischen Prognosemuster geben. Zugleich zeigen Erwartungstheorie und empirische Beobachtungen, dass Markterwartungen in bestimmten Regimen vorhersagbare Strukturen aufweisen können (Diebold, Rudebusch & Aruoba, 2006). Damit steht die Zinsprognose im Spannungsfeld zwischen theoretischer Effizienz und praktischer Vorhersagbarkeit, zwischen dem Ideal vollkommener Information und der Realität volatiler Erwartungen. Diese Arbeit untersucht, inwieweit sich Abweichungen von der Markteffizienz durch ökonometrische Modelle systematisch ausnutzen lassen. Sollten Modelle mit fundamentalen Variablen die naive Prognose signifikant übertreffen, wäre dies ein Indiz für adaptive oder begrenzt rationale Informationsverarbeitung am US-Zinsmarkt.

Vor diesem Hintergrund erweist sich die Prognose der 10-jährigen US-Treasury-Rendite als methodisch anspruchsvolle Aufgabe. Neben klassischen makroökonomischen Determinanten wie Inflationserwartungen, Bruttoinlandsprodukt-Wachstum, Staatsverschuldung und geldpolitischen Entscheidungen wirken zunehmend auch internationale Kapitalströme und technologische Veränderungen auf die Zinsbildung ein (Koroleva & Kopeykin, 2022).

Die empirische Untersuchung unterschiedlicher Modellklassen für die 10-jährige US-Treasury-Rendite gewinnt dadurch zunehmend an Bedeutung. Sie ermöglicht nicht nur eine Bewertung der statistischen Prognosekraft der Modelle, sondern auch eine Überprüfung ökonomischer Theorien zur Informationsverarbeitung und Markteffizienz im Zinsmarkt. Die vorliegende Arbeit verfolgt daher das Ziel, ökonomische Theorie und empirische Modellierung miteinander zu verknüpfen und zu prüfen, inwieweit sich die Erwartungstheorie der Zinsstruktur in einem weitgehend effizienten Markt empirisch abbilden lässt.

Konkret untersucht die Arbeit, welche ökonometrischen Modelle die höchste Prognosegüte für die 10-jährige US-Staatsanleihenrendite erzielen und welche ökonomischen Mechanismen dieser Prognosekraft zugrunde liegen. Damit wird nicht nur die statistische

Leistungsfähigkeit unterschiedlicher Modellklassen analysiert, sondern zugleich ein Beitrag zur Beantwortung einer grundlegenden Frage der Finanzmarktforschung geleistet. Sind Zinsmärkte tatsächlich so effizient, wie es die Theorie postuliert oder existieren beständige Erwartungsmuster, die eine zumindest partielle Vorhersagbarkeit ermöglichen?

## 2 Methodische und theoretische Grundlagen der Zinsprognose

Die theoretischen Grundlagen dieser Arbeit bilden das Fundament für die empirische Untersuchung der Prognosefähigkeit ökonometrischer Modelle im Kontext der 10-jährigen US-Staatsanleihenrendite. Die Ökonometrie, insbesondere die Zeitreihenökonomie, befasst sich dabei mit der Schätzung von Gleichungen, die stochastische Komponenten enthalten, um ökonomische Zusammenhänge quantitativ zu erfassen. Ziel des modernen Zeitreihenökonomikers ist es, Modelle zu entwickeln, die einerseits hinreichend einfach sind, um ökonomische Daten zuverlässig prognostizieren zu können, und andererseits eine fundierte Interpretation wirtschaftlicher Zusammenhänge sowie die Überprüfung theoretischer Hypothesen ermöglichen (Enders, 2015, ch. 1).

Wie bereits zuvor angedeutet, lassen sich zwei komplementäre Perspektiven unterscheiden. Einmal die Effiziente-Markt-Hypothese als theoretische Grundlage und die Erwartungstheorie der Zinsstruktur als ihr konzeptionelles Gegenstück. Während die Effiziente-Markt-Hypothese nach Fama (1970) davon ausgeht, dass Marktpreise sämtliche verfügbaren Informationen vollständig widerspiegeln und damit nicht systematisch prognostizierbar sind, interpretiert die Erwartungstheorie langfristige Zinssätze als das Ergebnis rational gebildeter Erwartungen über zukünftige kurzfristige Zinsen (Diebold, Rudebusch & Aruoba, 2006). Dieses Kapitel schafft damit die Grundlage, um zu untersuchen, in welchem Ausmaß Zinssätze als Ausdruck effizienter Informationsverarbeitung verstanden werden können, ob sich die in ihnen enthaltenen Erwartungen über zukünftige ökonomische Entwicklungen empirisch modellieren lassen und welche Modellansätze sich dabei als besonders geeignet erweisen.

### Die Effiziente-Markt-Hypothese (EMH)

Die Effiziente-Markt-Hypothese (Efficient Market Hypothesis, EMH) wurde durch Eugene Fama (1970) als eine der einflussreichsten Theorien der Finanzmarktforschung etabliert. Sie postuliert, dass alle verfügbaren Informationen, seien sie vergangenheitsbezogen, öffentlich oder sogar privat, zu jedem Zeitpunkt vollständig in den Marktpreisen reflektiert sind. In einem effizienten Markt folgt jede Preisänderung ausschließlich neuen Informationen, die per Definition zufällig und unvorhersehbar eintreten. Somit impliziert die EMH, dass die zukünftige Preisentwicklung, einschließlich der Entwicklung langfristiger Zinssätze, nicht systematisch prognostizierbar ist (Fama, 1970).

Für den Renten- und Zinsmarkt bedeutet dies, wenn der Markt effizient ist, können weder technische Analyse noch ökonometrische Prognosemodelle eine nachhaltige Informationsüberlegenheit erzielen. Jede Verbesserung der Prognosegüte wäre dann nur temporär und statistisch zufällig.

Gleichwohl hat die empirische Forschung in den letzten Jahrzehnten gezeigt, dass Finanzmärkte, insbesondere Zinsmärkte, nicht immer vollkommen effizient agieren.

Strukturelle Brüche, verzögerte Informationsverarbeitung, Verhaltensanomalien und geldpolitische Eingriffe können dazu führen, dass Erwartungen systematisch verzerrt sind (Shiller, 1980; Koroleva & Kopeykin, 2022). Diese Beobachtungen öffnen die theoretische Tür für Prognoseansätze, die davon ausgehen, dass sich in Marktbewegungen zumindest zeitweise deterministische Muster verbergen. Damit steht die EMH nicht im Widerspruch, sondern in einem Spannungsverhältnis zur Erwartungstheorie der Zinsstruktur, einer Theorie, die langfristige Zinsen explizit als aggregierte Erwartungsgröße interpretiert und damit eine potenzielle Grundlage für Prognosemodelle liefert.

### Die Erwartungstheorie der Zinsstruktur

Die Erwartungstheorie der Zinsstruktur (Expectations Hypothesis of the Term Structure, EHTS) bildet das zentrale theoretische Konzept zur Erklärung langfristiger Zinssätze. Nach der Erwartungshypothese ergibt sich der Zinssatz einer langfristigen Anleihe als arithmetisches Mittel der erwarteten zukünftigen kurzfristigen Zinssätze, ergänzt um eine Risikoprämie, die als Termprämie bezeichnet wird (Diebold, Rudebusch & Aruoba, 2006).

Formal lässt sich die Theorie wie folgt ausdrücken:

$$i_t^n = \left(\frac{1}{n}\right) \sum_{k=0}^{n-1} E_t(i_{t+k}^1) + TP_t^n$$

wobei  $i_t^n$  der nominale Zinssatz einer n-jährigen Anleihe,  $E_t(i_{t+k}^1)$  der erwartete zukünftige Ein-Jahres-Zins und  $TP_t^n$  die Termprämie ist.

Diese Darstellung zeigt, dass langfristige Zinssätze nicht nur aktuelle Marktbedingungen reflektieren, sondern vor allem die Erwartungen der Marktteilnehmer über die zukünftige geldpolitische und wirtschaftliche Entwicklung. Damit wird die Zinsstrukturkurve, das Verhältnis zwischen kurz- und langfristigen Zinssätzen, zu einem Indikator aggregierter Markterwartungen (Guidolin & Ionta, 2025).

Empirische Befunde belegen, dass die Form der Zinsstrukturkurve häufig Aufschluss über den makroökonomischen Zyklus gibt. Eine normale, ansteigende Kurve deutet auf Wachstumserwartungen hin, während eine inverse Kurve historisch mit Rezessionen korreliert (Estrella & Hardouvelis, 1991; Joshi, o.J.). Diese empirische Beziehung illustriert, dass sich Markterwartungen zwar nicht präzise, aber in ihrer Tendenz modellieren lassen.

### Die Fisher-Gleichung (Inflation–Zins–Beziehung)

Ein wesentliches Bindeglied zwischen Geldpolitik, Inflation und realwirtschaftlicher Aktivität stellt die sogenannte Fisher-Gleichung dar. Sie beschreibt die theoretische Beziehung zwischen dem nominalen Zinssatz, dem realen Zinssatz und der erwarteten Inflationsrate und bildet damit eine der grundlegenden Identitäten der modernen Makroökonomie (Fisher, 1930, ch. 2; Mishkin, 2019, ch. 4). Formal lässt sich diese Beziehung wie folgt ausdrücken:

$$i_t = r_t + \pi_t^e$$

wobei  $i_t$  den nominalen Zinssatz,  $r_t$  den realen Zinssatz und  $\pi_t^e$  die erwartete Inflationsrate zum Zeitpunkt  $t$  bezeichnet.

Die Fisher-Gleichung verdeutlicht, dass die nominalen Zinssätze langfristig von zwei Komponenten bestimmt werden. Von dem realen Ertrag auf Kapital und den Inflationserwartungen der Marktteilnehmer. Eine Erhöhung der erwarteten Inflationsrate führt, bei konstantem realem Zins, zu einem proportionalen Anstieg des nominalen Zinssatzes. In der geldpolitischen Praxis ist dieser Zusammenhang zentral, da er die Transmission geldpolitischer Impulse über Erwartungen vermittelt. Wird eine anhaltend höhere Inflation erwartet, so steigen die nominalen Renditen am Anleihemarkt, selbst wenn die kurzfristigen Leitzinsen noch unverändert bleiben (Mishkin, 2019, ch. 5). Für die Zinsprognose bedeutet dies, dass Inflationserwartungen, etwa gemessen über Breakeven-Inflationsraten oder Umfragen wie die „Survey of Professional Forecasters“, einen entscheidenden Beitrag zur Erklärung langfristiger Renditen leisten können (Wang et al., 2025). Empirisch zeigt sich jedoch häufig eine inkomplette Transmission von Inflationserwartungen in nominale Zinsen (Fisher-Puzzle). Dies könnte auf Geldillusion, Risikoprämien oder Zentralbankglaubwürdigkeit zurückzuführen sein (Cooray, 2002).

In Verbindung mit der Erwartungstheorie der Zinsstruktur ergänzt die Fisher-Gleichung das Verständnis der Zinsbildung um eine reale Dimension. Während die Erwartungstheorie den Zusammenhang zwischen kurz- und langfristigen nominalen Zinsen beschreibt, macht die Fisher-Gleichung deutlich, wie Inflationserwartungen in diesen Prozess eingebettet sind. Gemeinsam bilden beide Ansätze ein konsistentes theoretisches Fundament, das erklärt, wie Preis- und Zinserwartungen ineinandergreifen und sich in den Renditen langfristiger Staatsanleihen widerspiegeln.

Die empirische Analyse langfristiger Anleiherenditen erfordert auch ein grundlegendes Verständnis der Mechanismen, die Zinsbewegungen über verschiedene Zeithorizonte hinweg bestimmen. Während kurzfristige Schwankungen der Renditen häufig auf geldpolitische Entscheidungen, Marktstimmungen und Liquiditätsbedingungen reagieren, spiegeln langfristige Zinsen die aggregierten Erwartungen über Inflation, Wachstum und Risikoaufschläge wider (Diebold & Rudebusch, 2013, ch. 1; Guidolin & Ionta, 2025).

Diese Unterscheidung zwischen kurz- und langfristigen Dynamiken ist entscheidend für die Wahl geeigneter Prognosemodelle. Kurzfristige Modelle zielen darauf ab, kurzfristige Volatilität und Markterwartungen möglichst präzise abzubilden, während langfristig orientierte Ansätze versuchen, strukturelle makroökonomische Zusammenhänge zu erfassen (Carnot et al., 2005 ch. 5). In der Literatur hat sich daher eine klare Trennung zwischen zeitreihenbasierten Modellen, die ausschließlich historische Informationen der Rendite selbst nutzen, und strukturellen oder multivariaten Modellen, die externe ökonomische Einflussgrößen einbeziehen, etabliert (Lütkepohl, 2005, 10.4).

Für die vorliegende Arbeit ist diese Differenzierung sehr relevant, da die 10-jährige Treasury-Rendite sowohl von kurzfristigen geldpolitischen Maßnahmen der Federal

Reserve als auch von langfristigen Erwartungen über Inflation, Staatsverschuldung und Wachstum beeinflusst wird. Die folgenden Abschnitte führen daher zunächst in die theoretischen und statistischen Grundlagen der Zeitreihenanalyse ein, bevor verschiedene Modellklassen, univariate, multivariate und fehlerkorrigierende Ansätze, vorgestellt werden, die darauf abzielen, die kurzfristigen und langfristigen Schwankungen mit einzubeziehen und dann die Gleichgewichtsbeziehungen zu überbrücken.

## 2.1 Grundlagen der Zeitreihenanalyse

Zeitreihenanalysen bilden das methodische Fundament nahezu aller empirischen Untersuchungen zu Finanz- und Zinsmärkten. Eine Zeitreihe beschreibt den numerischen Verlauf einer ökonomischen Größe über die Zeit, wobei jede Beobachtung  $y_t$  nicht isoliert, sondern in Abhängigkeit von vergangenen Werten betrachtet wird (Box et al., 2016, ch. 1). Diese Eigenschaft, die als Autokorrelation bezeichnet wird, beschreibt die systematische Abhängigkeit gegenwärtiger Beobachtungen von ihren eigenen Vergangenheitswerten und bildet damit die Grundlage für die Erkennung zeitlicher Strukturen sowie die Entwicklung belastbarer Prognosemodelle.

Da ökonomische Reihen häufig Trend- oder Saisonalstrukturen enthalten, müssen sie vor der Modellierung überprüft und gegebenenfalls durch Transformationen wie Logarithmierung oder Differenzierung stationär gemacht werden (Hyndman & Athanasopoulos, 2021, ch. 8.1). Nur stationäre Reihen erlauben eine stabile Parameterschätzung und eine sinnvolle Interpretation zukünftiger Entwicklungen. Entsprechend kommen Tests wie der Augmented Dickey-Fuller-Test (ADF) zum Einsatz, um das Vorliegen von Einheitswurzeln und die Stabilität der Varianz zu prüfen (Box et al. 2016, ch. 10).

Damit liefert die Zeitreihenanalyse nicht nur den technischen Rahmen für die Modellierung, sondern bildet auch den empirischen Zugang zu einem Markt, der sich theoretisch der Prognose entziehen sollte. Sie steht damit sinnbildlich für das Kernanliegen dieser Arbeit, die Verbindung von theoretischer Markteffizienz und praktischer Vorhersagbarkeit.

### Stationarität, stochastische Prozesse und Transformationen

Ein zentrales Konzept der Zeitreihenanalyse ist die Stationarität. Eine Zeitreihe gilt als stationär, wenn ihre statistischen Eigenschaften, wie der Mittelwert, die Varianz und Kovarianz, über die Zeit hinweg konstant bleiben. Stationarität ist eine grundlegende Voraussetzung für viele Prognosemodelle, da sie sicherstellt, dass die Schätzparameter zeitinvariant bleiben und somit stabile Vorhersagen ermöglichen (Box et al. 2016, ch. 2).

In der Praxis weisen jedoch die meisten makroökonomischen Zeitreihen, wie Zinsen, Trend- oder Saisoneffekte auf und sind damit nicht stationär. Solche Reihen werden in der Ökonometrie häufig als stochastische Prozesse verstanden, bei denen jede Beobachtung eine Realisierung einer Zufallsvariablen darstellt. Eine besondere Bedeutung kommt hierbei Prozessen mit einer Einheitswurzel zu, da diese auf Nicht-Stationarität hinweisen. Ein

Einheitswurzelprozess, beispielsweise ein Random Walk, besitzt einen persistierenden Trend, der durch Schocks dauerhaft beeinflusst wird (Said et al., 1984).

Zur Überprüfung der Stationarität werden Einheitswurzeltests wie der Augmented-Dickey-Fuller-Test (ADF) eingesetzt (Dickey & Fuller, 1979). Dabei wird die Nullhypothese geprüft, dass eine Zeitreihe eine Einheitswurzel enthält. Wird diese Hypothese verworfen ( $p < 0,05$ ), gilt die Reihe als stationär. Bleibt sie bestehen, ist eine Transformation, meist durch Differenzierung, erforderlich, um die Voraussetzungen für eine Modellschätzung und Prognose zu erfüllen (Shumway & Stoffer, 2018, ch. 5.2). Zur Sicherstellung der Vergleichbarkeit und Robustheit der Modellierung wurden alle Variablen auf Stationarität geprüft. Die Beurteilung erfolgte anhand des Augmented-Dickey-Fuller-Tests (ADF). In einzelnen Fällen (z. B. FEDFUNDS, UNRATE) wurde jedoch trotz formaler Stationarität ( $p < 0,05$ ) eine zusätzliche Differenzierung vorgenommen, da visuell persistente Trendkomponenten und strukturelle Regimewechsel erkennbar waren. Dieses Vorgehen entspricht der in der Zeitreihenökonomie gängigen Praxis, bei makroökonomischen Variablen mit deutlicher Trendlage eine konservative Transformation vorzunehmen, um Scheinkorrelationen und instabile Parameterschätzungen zu vermeiden (Thome, 1994).

Die Differenzierung eliminiert deterministische Trends, sodass die verbleibende Reihe um einen konstanten Mittelwert schwankt und eine stabilere Varianz aufweist. Ergänzend werden, insbesondere bei Finanzmarktdaten mit starken Skalenunterschieden oder heteroskedastischen Strukturen, auch logarithmische Transformationen oder Mittelwertbereinigungen eingesetzt, um die Varianz zu stabilisieren und die Vergleichbarkeit über Zeiträume hinweg zu verbessern (Box et al., 2016, ch. 4).

### **Differenzierte Renditen**

Um die Stationarität der Zeitreihe herzustellen und strukturelle Trends zu eliminieren, werden in dieser Arbeit differenzierte Renditen verwendet. Dabei wird nicht das Zinsniveau selbst, sondern die Veränderung zwischen zwei aufeinanderfolgenden Beobachtungen analysiert. Die Differenzierung ermöglicht es, Wachstumsraten beziehungsweise Veränderungen der Rendite über die Zeit abzubilden und damit die dynamischen Anpassungsprozesse der Zinsentwicklung zu untersuchen (Tsay, 2010, ch. 1.1).

Formal lässt sich die erste Differenz einer Zeitreihe  $y_t$  wie folgt darstellen:

$$r_t = y_t - y_{t-1}$$

Hierbei beschreibt  $r_t$  die Veränderung der Rendite von einem Zeitpunkt zum nächsten. Durch die Differenzierung werden langfristige Trends entfernt, wodurch die resultierende Zeitreihe eine stabilere Varianz aufweist und besser für ökonometrische Modellierungen geeignet ist. In der vorliegenden Arbeit dienen daher differenzierte Werte der 10-jährigen US-Treasury-Rendite als Grundlage, um relative Bewegungen und kurzfristige Dynamiken anstelle absoluter Zinsniveaus zu analysieren.

## **Autokorrelation und Modellierung zeitlicher Abhängigkeiten**

Ein wesentliches Merkmal von Zeitreihen ist die Autokorrelation, also die Abhängigkeit eines aktuellen Wertes von seinen eigenen vergangenen Werten. Diese Eigenschaft unterscheidet Zeitreihen fundamental von Querschnittsdaten. Um die Struktur solcher Abhängigkeiten zu identifizieren, werden die Autokorrelationsfunktion (ACF) und die partielle Autokorrelationsfunktion (PACF) herangezogen. Diese dienen als zentrale Werkzeuge zur Voranalyse von Zeitreihendaten und zur Spezifikation geeigneter Modellordnungen (Box et al., 2016, ch. 6.2).

Die ACF misst die lineare Korrelation zwischen der Zeitreihe und ihren eigenen Vergangenheitswerten fester Lags. Die PACF hingegen gibt die Korrelation zwischen der Serie und ihren verzögerten Werten an, nachdem der Einfluss zwischenliegender Beobachtungen herausgerechnet wurde (Box et al., 2016, ch. 6.2).

### **Mittelwertbereinigung (Demeaning)**

Zur weiteren Stabilisierung wird die Zeitreihe mittelwertbereinigt, indem der arithmetische Mittelwert subtrahiert wird:

$$x_t^* = x_t - \bar{x}$$

Dadurch wird die Reihe um ihren langfristigen Durchschnittszins zentriert, was insbesondere für autoregressive Modelle wichtig ist, da es den konstanten Term eliminiert und die Interpretation der Modellkoeffizienten vereinfacht (Armstrong, 2001, ch. 4).

### **Rücktransformation**

Nach der Modellschätzung erfolgt die Rücktransformation der Prognosen in die ursprüngliche Skala der Zeitreihe. Hierzu wird der zuvor subtrahierte Mittelwert wieder addiert:

$$\hat{y}_t = \hat{y}_t^* + \bar{x}$$

Dieser Schritt ist notwendig, um die Ergebnisse ökonomisch interpretierbar zu machen (Armstrong, 2001, ch. 4).

## **2.2 Univariate Prognosemodelle: Die ARIMA-Klasse**

Die im vorangegangenen Kapitel eingeführten Grundlagen der Zeitreihenanalyse bilden die methodische Basis für die Anwendung univariater Prognosemodelle. Zu den bekanntesten und am weitesten verbreiteten Ansätzen zählt das ARIMA-Modell (AutoRegressive Integrated Moving Average), das seit den Arbeiten von Box und Jenkins (1976) als Standardverfahren in der ökonometrischen Prognose gilt. Es kombiniert autoregressive (AR) und gleitende Durchschnittskomponenten (MA) und erweitert diese um eine Integrationskomponente (I), welche durch Differenzierung nicht-stationäre Zeitreihen in stationäre Prozesse überführt.

Das Ziel eines ARIMA-Modells besteht darin, zeitliche Abhängigkeiten innerhalb einer einzelnen Variablen, hier der Rendite zehnjähriger US-Treasuries zu identifizieren und für Prognosezwecke nutzbar zu machen. Der Ansatz beruht auf der Annahme, dass gegenwärtige Werte einer Zeitreihe  $y_t$  wesentlich durch ihre eigenen Vergangenheitswerte sowie durch stochastische Störungen beeinflusst werden.

Ein ARIMA( $p, d, q$ )-Prozess ist eine Erweiterung des klassischen ARMA( $p, q$ )-Prozesses und lässt sich mithilfe des Lag-Operators  $L$  kompakt darstellen als (Box et al., 2016, ch. 4):

$$\Phi(L)(1 - L)^d y_t = \theta(L)\varepsilon_t$$

Wobei  $\Phi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p$  das autoregressive Lag-Polynom der Ordnung  $p$  bezeichnet,  $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$  das Moving-Average-Lag-Polynom der Ordnung  $q$  ist,  $d$  die Anzahl der Differenzierungen angibt, die zur Herstellung von Stationarität notwendig sind, und  $\varepsilon_t$  einen weißen Rauschprozess mit Erwartungswert null und Varianz  $\sigma_\varepsilon^2$  beschreibt.

Das Modell wurde ursprünglich im Rahmen der von Box und Jenkins (1976) entwickelten Methodologie konzipiert und bildet bis heute eine der zentralen Grundlagen der Zeitreihenanalyse (Neusser, 2025, ch. 2).

Die Parameter eines ARIMA-Modells werden üblicherweise mit der Maximum-Likelihood-Methode (ML) geschätzt, da sie eine effiziente und konsistente Schätzung der Modellkoeffizienten ermöglicht. Ziel ist es, die Werte der Parameter zu finden, welche die Likelihood-Funktion, also die Wahrscheinlichkeit der beobachteten Daten unter dem Modell, maximieren.

Zur Auswahl des bestgeeigneten Modells wird anschließend das Akaike Information Criterion (AIC) herangezogen, das einen Kompromiss zwischen Modellgüte und Komplexität herstellt (Enders, 2015, ch. 2). Das AIC wird definiert als

$$AIC = -2 \ln(L) + 2k$$

wobei  $L$  den Wert der maximierten Likelihood-Funktion und  $k$  die Anzahl der geschätzten Parameter bezeichnet. Das Modell mit dem niedrigsten AIC-Wert wird bevorzugt, da es die beste Balance zwischen Anpassungsqualität und Parsimonie bietet (Neusser, 2025, ch. 11.4.2).

Durch die Kombination dieser drei Komponenten kann das Modell unterschiedliche Formen von Abhängigkeiten abbilden, kurzfristige Persistenzen, gleitende Fehlerkorrekturen oder längerfristige Trends (Box et al., 2016, ch 4.).

Das ARIMA-Modell ist ein datengetriebener Ansatz, der ausschließlich auf der internen Dynamik der Zeitreihe basiert und keine zusätzlichen erklärenden Variablen benötigt. Diese Eigenschaft macht es besonders geeignet für Märkte, in denen, entsprechend der Effizienten-Markt-Hypothese nach Fama (1970), alle relevanten Informationen bereits in den Preisen enthalten sind. In diesem Sinne dient das ARIMA-Modell auch als empirischer

Prüfstein dieser Annahme. Wenn Märkte effizient sind, sollte sich aus vergangenen Renditen keine systematische Prognose zukünftiger Bewegungen ableiten lassen.

Gleichzeitig ist der Ansatz theoretisch begrenzt, da er keine makroökonomischen Einflussgrößen berücksichtigt. ARIMA-Modelle liefern damit zwar statistisch konsistente, aber ökonomisch schwer interpretierbare Ergebnisse. Sie bilden daher die Ausgangsbasis der empirischen Analyse und dienen als Benchmark, um zu prüfen, ob die Einbindung exogener Variablen, wie Inflation, Wirtschaftswachstum oder Leitzinsen, in multivariaten Modellen einen zusätzlichen Informationsgewinn und damit eine höhere Prognosegüte ermöglicht.

### **2.3 Multivariate Prognosemodelle: Lineare Regression, ARIMAX und Fehlerkorrekturmodelle (ECM)**

Während univariate Modelle wie die ARIMA-Klasse ausschließlich vergangene Werte der Zielvariablen berücksichtigen, erweitern multivariate Prognosemodelle den analytischen Horizont um exogene Einflussgrößen. Diese Modelle beruhen auf der Annahme, dass die Dynamik ökonomischer Variablen, insbesondere von Zinsrenditen, nicht nur durch ihre eigene Historie, sondern auch durch fundamentale makroökonomische Faktoren bestimmt wird.

Die Einbindung erklärender Variablen folgt der Annahme, dass langfristige Zinsen das Ergebnis von Erwartungen über zukünftige wirtschaftliche Entwicklungen sind (Diebold, Rudebusch & Aruoba, 2006). Im Sinne der Erwartungstheorie spiegeln 10-jährige Renditen daher nicht nur erwartete Kurzfristzinsen, sondern auch Informationen über Inflation, Wachstum oder geldpolitische Maßnahmen wider. Multivariate Modelle ermöglichen es, diese theoretischen Zusammenhänge empirisch zu quantifizieren und ihren Beitrag zur Prognosegüte zu prüfen. Aus Sicht der Effizienten-Markt-Hypothese (Fama, 1970) ist diese Erweiterung besonders aufschlussreich, denn führt der Einbezug makroökonomischer Variablen zu einer signifikanten Verbesserung der Prognose, deutet dies auf temporäre Ineffizienzen oder Verzögerungen in der Informationsverarbeitung hin.

#### **Lineare Regressionsmodelle**

Das lineare Regressionsmodell stellt die einfachste Form eines multivariaten Prognoseansatzes dar. Es beschreibt die abhängige Variable, beispielsweise die Rendite zehnjähriger US-Treasuries, als lineare Funktion mehrerer unabhängiger Variablen, die deren Entwicklung theoretisch beeinflussen.

Die Grundform lautet:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \dots + \beta_k x_{k,t} + \varepsilon_t$$

Hierbei steht  $y_t$  für die abhängige Variable,  $x_{(i,t)}$  für die unabhängigen Variablen und  $\varepsilon_t$  für den Störterm, der unbeobachtete Einflüsse enthält (Greene, 2002, ch. 2.2). Die Parameter  $\beta_i$  werden üblicherweise mit der Methode der kleinsten Quadrate (*Ordinary Least Squares*,

OLS) geschätzt, welche die Summe der quadrierten Abweichungen zwischen beobachteten und vorhergesagten Werten minimiert (Greene, 2002, ch. 3.2.1).

In der Zinsanalyse kann dieses Modell beispielsweise genutzt werden, um den Einfluss von Inflationsraten, Wirtschaftswachstum, Arbeitsmarktdaten oder Aktienmarktdaten auf die Renditeentwicklung zu quantifizieren. Der Vorteil liegt in der hohen Interpretierbarkeit und Transparenz der Ergebnisse. Allerdings unterstellt das Modell lineare Zusammenhänge und zeitlich konstante Parameter, was in volatilen Finanzmärkten oft nicht vollständig realistisch ist (Hyndman & Athanasopoulos, 2021, ch. 5.8). Daher dient die klassische lineare Regression häufig als Ausgangspunkt für komplexere dynamische Modelle.

### Das ARIMAX-Modell

Das ARIMAX-Modell (*AutoRegressive Integrated Moving Average with eXogenous Variables*) erweitert das ARIMA-Modell um erklärende Variablen. Es kombiniert die internen Zeitreihenstrukturen (autoregressiv und gleitender Durchschnitt) mit externen Einflussfaktoren.

Formal lässt sich das Modell wie folgt darstellen:

$$y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + n_t,$$

wobei  $n_t$  den stochastischen Anteil beschreibt, der selbst einem ARIMA( $p, d, q$ )-Prozess folgt:

$$\Phi(B)(1 - B)^d n_t = \Theta(B)w_t, w_t \sim WN(0, \sigma^2).$$

Dadurch quantifizieren die Regressionskoeffizienten  $\beta_i$  den direkten Einfluss der exogenen Variablen, während die ARIMA-Struktur im Fehlerterm verbleibende zeitliche Abhängigkeiten modelliert (Shumway & Stoffer, 2018., ch. 5.6).

Das ARIMAX-Modell bietet den Vorteil, dass es sowohl die historische Abhängigkeit der Zielgröße als auch den Einfluss externer Faktoren abbildet. In der Zinsprognose lassen sich damit geldpolitische Impulse, makroökonomische Variablen oder Stimmungsindikatoren berücksichtigen.

### Kointegration und Error-Correction-Modelle (ECM)

Die zuvor vorgestellten Modelle (ARIMA und ARIMAX) eignen sich primär zur Modellierung kurzfristiger Dynamiken und stationärer Zeitreihen. In der Praxis weisen ökonomische Variablen, insbesondere Zinssätze, Preisniveaus oder Outputgrößen, jedoch häufig einen stochastischen Trend auf. Werden solche nicht-stationären Reihen direkt regressiert, besteht die Gefahr einer Scheinkorrelation (spurious regression): Das Modell suggeriert einen Zusammenhang, der rein zufällig auf gemeinsamen Trends beruht (Granger & Newbold, 1974). Um diesem Problem zu begegnen, wurde das Konzept der Kointegration entwickelt, das es erlaubt, langfristige Gleichgewichtsbeziehungen zwischen nicht-stationären Variablen zu modellieren (Engle & Granger, 1987).

### Konzept der Kointegration und langfristige Gleichgewichtsbeziehungen

Zwei oder mehr Zeitreihen gelten als kointegriert, wenn eine lineare Kombination dieser Variablen stationär ist, obwohl die Einzelreihen selbst einen stochastischen Trend aufweisen. Formal gilt für die zwei integrierte Prozesse  $x_t$  und  $y_t$  der Ordnung  $I(1)$ :

$$y_t = \alpha + \beta x_t + u_t, u_t \sim I(0)$$

Der Störterm  $u_t$  repräsentiert die kurzfristigen Abweichungen vom langfristigen Gleichgewicht (Das, 2023, ch. 12.3 – 12.5). Kointegration impliziert somit, dass ökonomische Variablen über längere Zeiträume durch fundamentale Gleichgewichtsbeziehungen miteinander verbunden sind, auch wenn sie kurzfristig voneinander abweichen (Lütkepohl, 2005, ch. 6.3).

### Die Engle-Granger-Methode (Zwei-Schritt-Verfahren)

Ein etabliertes Verfahren zur empirischen Überprüfung auf Kointegration ist die Engle-Granger-Methode (Engle & Granger, 1987). Sie besteht aus zwei Schritten: Im ersten Schritt wird eine langfristige Gleichgewichtsbeziehung zwischen den nicht-stationären Variablen geschätzt, beispielsweise:

$$y_t = \alpha + \beta x_t + u_t$$

Im zweiten Schritt werden die geschätzten Residuen  $\hat{u}_t$  auf Stationarität geprüft, meist mithilfe eines erweiterten Dickey-Fuller-Tests (ADF). Wird die Nullhypothese einer Einheitswurzel verworfen, gelten die Variablen als kointegriert. Dieses Verfahren bildet die Grundlage für die Spezifikation eines Error-Correction-Modells (ECM), das kurzfristige Anpassungen an langfristige Gleichgewichte beschreibt.

### Das Error-Correction-Modell (ECM)

Das ECM verknüpft kurzfristige Dynamiken mit einem Mechanismus, der langfristige Gleichgewichte wiederherstellt. Formal lässt sich das ECM als Transformation eines autoregressiven verteilten Lag-Modells (ADL) darstellen. Für zwei Variablen  $x_{1t}$  und  $x_{2t}$  ergibt sich die Fehlerkorrekturform:

$$\Delta x_{1t} = \alpha_0 + \beta_1 \Delta x_{2t} - \lambda(x_{1,t-1} - \gamma x_{2,t-1}) + u_t$$

Der Term  $(x_{1,t-1} - \gamma x_{2,t-1})$  stellt das Ungleichgewicht der vergangenen Periode dar und wird als Fehlerkorrekturterm bezeichnet. Der Koeffizient  $\lambda$  beschreibt die Anpassungsgeschwindigkeit. Ein negativer und signifikanter Wert bedeutet, dass eine positive Abweichung vom Gleichgewicht zu einer gegenläufigen Korrektur der abhängigen Variable führt. Dadurch wird das langfristige Gleichgewicht schrittweise wiederhergestellt.

Das ECM kombiniert somit kurzfristige Veränderungen ( $\Delta x_{2t}$ ) mit einem Mechanismus, der das System auf seinen langfristigen Gleichgewichtspfad zurückführt (Das, 2023, ch. 12.3 – 12.5).

Für die Prognose langfristiger Anleiherenditen, wie der 10-jährigen US-Treasury-Note, ist dieses Modell besonders relevant. Es erlaubt, die kurzfristigen Marktreaktionen auf geldpolitische Entscheidungen mit den langfristigen Fundamentalfaktoren zu verknüpfen. So lassen sich Abweichungen der Rendite von ihrem theoretischen Gleichgewicht quantitativ erfassen und in zukünftige Entwicklungen überführen. Ein signifikanter Korrekturterm würde darauf hindeuten, dass langfristige Gleichgewichte bestehen und Märkte nicht vollständig effizient sind (Lütkepohl, 2005, ch. 6.3).

## 2.4 Gütekriterien zur Evaluierung von Prognosen

Die Bewertung der Prognosequalität stellt einen zentralen Bestandteil ökonometrischer Analysen dar. Da unterschiedliche Modelle, von einfachen linearen Regressionen bis hin zu komplexen ARIMAX- oder ECM-Strukturen, verwendet werden, ist ein objektiver Vergleich ihrer Vorhersagegüte essenziell. Gütekriterien ermöglichen es, die Genauigkeit, Robustheit und praktische Relevanz der erstellten Prognosen zu beurteilen (Vo & Ślepaczuk, 2022).

### Fehlerbasierte Gütemaße

Fehlerbasierte Gütemaße quantifizieren die Abweichung zwischen den prognostizierten und den tatsächlich beobachteten Werten einer Zeitreihe. Sie bilden die Grundlage zur Beurteilung der Modellgüte und ermöglichen den Vergleich alternativer Prognoseansätze auf objektiver Basis (Hyndman & Athanasopoulos, 2021, ch. 3.4). Zu den wichtigsten Metriken zählen der Mean Absolute Error (MAE), der Mean Squared Error (MSE) und die Korrelation (COR) zwischen Prognose und Beobachtung.

### Mean Absolute Error (MAE)

$$MAE = \left(\frac{1}{T}\right) \sum_{t=1}^T |y_t - \hat{y}_t|$$

Der Mean Absolute Error (MAE) misst die durchschnittliche absolute Abweichung zwischen den prognostizierten Werten  $\hat{y}_t$  und den tatsächlichen Beobachtungen  $y_t$ . Ein niedriger MAE-Wert weist auf eine hohe Prognosegenauigkeit hin.

Da alle Abweichungen gleich stark gewichtet werden, eignet sich der MAE besonders für Datensätze mit moderater Streuung oder vereinzelt Ausreißern (Hyndman & Koehler, 2006, Vo & Ślepaczuk, 2022).

### Mean Squared Error (MSE)

$$MSE = \left(\frac{1}{T}\right) \sum_{t=1}^T (y_t - \hat{y}_t)^2$$

Der Mean Squared Error berücksichtigt, im Gegensatz zum MAE, die quadrierte Abweichung und gewichtet dadurch größere Fehler überproportional. Dadurch reagiert der MSE sensibler auf starke Prognoseabweichungen und ist besonders geeignet, um Modelle

mit potenziellen Strukturbrüchen oder Regimewechseln zu identifizieren (Hyndman & Koehler, 2006, Vo & Ślepaczuk, 2022).

Da große Fehler stärker gewichtet werden, stellt der MSE ein aussagekräftiges Maß für die Stabilität eines Modells dar. Ein niedriger MSE-Wert weist darauf hin, dass das Modell nicht nur im Durchschnitt, sondern auch bei extremen Beobachtungen präzise arbeitet. Ziel einer guten Prognose ist es daher, sowohl den MAE als auch den MSE zu minimieren, um eine möglichst hohe Gesamtgenauigkeit zu erreichen.

### **Korrelation (COR)**

$$COR = \frac{\text{Cov}(y_t, \hat{y}_t)}{(\sigma_y \cdot \sigma_{\hat{y}})}$$

Neben der Höhe der Fehler ist auch die Richtungstreue der Prognose von Bedeutung. Die Korrelation (COR) zwischen tatsächlichen und prognostizierten Werten misst, wie stark beide Reihen miteinander kovariieren (Stock, 2020, ch. 3.7).

Ein Wert von  $COR = 1$  bedeutet perfekte positive Korrelation (Prognose und Realität verlaufen identisch),  $COR = 0$  weist auf keinen Zusammenhang hin, und  $COR = -1$  steht für eine perfekte Gegenläufigkeit (Hoshmand, 2010, ch. 7.3).

In der Zinsprognose ist eine hohe Korrelation besonders wünschenswert, da sie anzeigt, dass das Modell nicht nur das Niveau, sondern auch die Bewegungsrichtung der Renditen korrekt erfasst (Paraskevopoulos & Santos, 2002).

Während MAE und MSE den Fehler im Niveau quantifizieren, beschreibt die Korrelation somit die Prognosequalität im Trendverlauf, ein wichtiger Aspekt, wenn Modelle zur Ableitung von Handelsstrategien eingesetzt werden. Die kombinierte Betrachtung dieser Kennzahlen liefert ein umfassendes Bild der Prognosequalität. MAE bewertet die durchschnittliche Abweichung in Originaleinheiten, MSE betont die Robustheit gegenüber Ausreißern und die Korrelation misst die Richtungstreue. Ein Modell mit geringem MAE, niedrigem MSE und hoher Korrelation wird als prognosestark und stabil eingestuft. Umgekehrt deuten hohe Fehlerwerte oder geringe Korrelationen darauf hin, dass das Modell systematische Muster in der Zeitreihe nicht adäquat erfasst.

### **Richtungsbasierte Gütemaße**

In der Finanzmarktanalyse ist nicht nur die Größe des Prognosefehlers relevant, sondern auch, ob die Richtung der vorhergesagten Veränderung korrekt ist. In Handelsanwendungen ist nicht die Präzision der Punktprognose ausschlaggebend, sondern die Fähigkeit des Modells, die Bewegungsrichtung der Rendite korrekt zu antizipieren. Zur Bewertung dieser Richtungstreue werden richtungsorientierte Performanzmaße herangezogen, insbesondere die Hit Ratio (HR) und die Weighted Hit Ratio (WHR).

**Hit Ratio (HR)**

$$HR = \left(\frac{1}{T}\right) \sum_{t=1}^T I[(y_t - y_{t-1})(\hat{y}_t - y_{t-1}) > 0]$$

Die Hit Ratio (HR) misst den Anteil der Zeitpunkte, an denen das Modell die Richtung der Renditeveränderung korrekt vorhergesagt hat. Der Indikator I nimmt den Wert 1 an, wenn die prognostizierte Veränderung  $(\hat{y}_t - y_{t-1})$  dieselbe Richtung aufweist wie die tatsächliche Veränderung  $(y_t - y_{t-1})$ , andernfalls nimmt der Indikator den Wert 0 an (Pesaran & Timmermann, 1992).

Ein Wert von 0,5 entspricht einem Zufallsergebnis, das heißt, das Modell ist im Mittel nicht besser als eine zufällige Schätzung. Werte über 0,5 deuten dagegen auf eine über den Zufall hinausgehende Prognosefähigkeit hin. Die Hit Ratio eignet sich insbesondere zur Beurteilung der ökonomischen Relevanz eines Modells, da sie unmittelbar auf die Entscheidungslogik von Marktteilnehmenden, insbesondere Kauf- und Verkaufsentscheidungen, übertragbar ist.

**Weighted Hit Ratio (WHR)**

$$WHR = \frac{[\sum_{t=1}^T |y_t - y_{t-1}| \cdot I[(y_t - y_{t-1})(\hat{y}_t - y_{t-1}) > 0]]}{[\sum_{t=1}^T |y_t - y_{t-1}|]}$$

Die Weighted Hit Ratio (WHR) erweitert die HR um eine Gewichtung nach der ökonomischen Bedeutung der jeweiligen Marktbewegungen. Während die HR alle korrekten Vorhersagen gleich behandelt, berücksichtigt die WHR die Größe der jeweiligen Veränderung. Große Marktbewegungen, etwa starke Zinsanstiege oder -rückgänge, erhalten somit ein höheres Gewicht als geringfügige Schwankungen (Gneiting & Raftery, 2007).

Ein WHR-Wert von 0,7 bedeutet beispielsweise, dass das Modell nicht nur häufig die richtige Richtung prognostiziert, sondern vor allem bei größeren Marktbewegungen überdurchschnittlich oft korrekt lag. Dadurch misst die WHR die wirtschaftliche Güte einer Prognose, nicht nur ihre statistische Präzision. Diese Kennzahl ist insbesondere im Kontext von Handelsstrategien oder Risikomanagementsystemen von Bedeutung, da dort die korrekte Einschätzung großer Bewegungen maßgeblich über Gewinn oder Verlust entscheidet.

**Risiko- und Performanz Kennzahlen**

Zur praktischen Beurteilung von Prognosen, die in Handelsstrategien umgesetzt werden, reicht die rein statistische Analyse der Prognosefehler oft nicht aus. Entscheidend ist vielmehr, ob ein Modell ökonomisch verwertbare Ergebnisse generiert, das heißt, ob die daraus abgeleiteten Signale zu einer verbesserten risikoadjustierten Rendite führen. Dafür

werden ergänzend finanzwirtschaftliche Performanz- und Risikokennzahlen herangezogen, die die Umsetzung der Prognosen in reale Marktentscheidungen abbilden.

### **Kumulative Rendite**

Die kumulative Rendite misst die Gesamtperformance eines Portfolios über den gesamten Beobachtungszeitraum und zeigt, um welchen Faktor der ursprüngliche Wert gestiegen oder gefallen ist. Sie berechnet sich aus dem Verhältnis von End- zu Anfangswert:

$$R_{\text{cum}} = \left( \frac{V_T}{V_0} \right) - 1$$

wobei  $V_T$  den Endwert und  $V_0$  den Anfangswert des Portfolios darstellt. Die kumulative Rendite gibt an, welchen Gesamtertrag eine Anlagestrategie über den betrachteten Zeitraum erzielt hat, ohne diesen auf Jahresbasis zu normieren. Sie eignet sich daher besonders, um die absolute Wertentwicklung einer Strategie zu bewerten (Sharpe, 1966; Bodie et al., 2018, ch. 5).

### **Annualisierte Rendite**

Die annualisierte Rendite misst die durchschnittliche jährliche Rendite über den gesamten Beobachtungszeitraum. Sie wird berechnet, indem die kumulierte Gesamtrendite auf ein Jahr hochgerechnet wird:

$$R_{\text{annual}} = \left( \frac{V_T}{V_0} \right)^{\frac{1}{n}} - 1$$

wobei  $V_T$  und  $V_0$  den End- bzw. Anfangswert des Portfolios und  $n$  die Anzahl der Jahre bezeichnen.

Die annualisierte Rendite erlaubt den Vergleich unterschiedlicher Strategien oder Zeiträume auf einer einheitlichen Basis. Ein höherer Wert signalisiert eine bessere durchschnittliche Performance pro Jahr. Allerdings berücksichtigt diese Kennzahl kein Risiko, ein Modell kann also eine hohe Rendite erzielen, aber gleichzeitig stark schwanken. Deshalb wird sie häufig zusammen mit der Volatilität und der Sharpe Ratio interpretiert (Sharpe, 1966; Bodie et al., 2018, ch. 5).

### **Volatilität**

Die Volatilität beschreibt die Schwankungsintensität der Renditen und dient als klassisches Risikomaß:

$$\sigma_p = \sqrt{\left[ \left( \frac{1}{T-1} \right) \sum_{t=1}^T (R_t - \bar{R})^2 \right]}$$

Eine hohe Volatilität zeigt an, dass die Renditen stark variieren, das bedeutet ein höheres Risiko, aber auch potenziell größere Chancen.

In der Bewertung von Prognosemodellen signalisiert eine niedrige Volatilität bei gleichbleibender Rendite eine stabilere Performance, während eine hohe Volatilität auf ein schwankungsanfälliges Modell hinweist (Markowitz, 1952).

Für Handelsstrategien ist daher nicht nur die absolute Rendite, sondern das Verhältnis von Rendite zu Risiko entscheidend, ein Gedanke, der in der Sharpe Ratio formalisiert wird.

### Sharpe Ratio

$$\text{Sharpe Ratio} = \frac{E[R_p - R_f]}{\sigma_p}$$

Die Sharpe Ratio misst die risikoadjustierte Rendite eines Portfolios. Sie setzt die durchschnittliche Überschussrendite ( $R_p - R_f$ ) gegenüber einer risikofreien Anlage ins Verhältnis zur Volatilität  $\sigma_p$ .

Ein hoher Sharpe-Wert zeigt, dass das Modell pro Risikoeinheit überdurchschnittlich viel Rendite erzielt. In der Literatur und Praxis gilt dabei die Faustregel, dass Werte über 1,0 als gut und Werte über 2,0 als sehr gut interpretiert werden, während Werte unter 1,0 auf eine nur geringe risikobereinigte Attraktivität hinweisen (Sharpe, 1994). Diese Schwellen sind nicht normativ festgelegt, haben sich jedoch als gängige Orientierung in der empirischen Finanzanalyse etabliert.

Für die Modellbewertung in dieser Arbeit lässt sich daraus ableiten, dass eine höhere Sharpe Ratio darauf hinweist, dass das Prognosemodell nicht lediglich statistische Präzision, sondern zugleich ökonomische Effizienz aufweist, indem es Informationsvorteile in tatsächlich erzielbare, risikoadjustierte Erträge transformiert.

### Maximum Drawdown (MDD)

Der Maximum Drawdown (MDD) misst den größten kumulierten Verlust eines Portfolios innerhalb des Untersuchungszeitraums und beschreibt somit das Abwärtsrisiko:

$$\text{MDD} = \max_{t \in [0, T]} \left[ \frac{P_{\text{peak}} - P_t}{P_{\text{peak}}} \right]$$

wobei  $P_{\text{peak}}$  den jeweils höchsten bisherigen Portfolio- oder Kapitalstand bezeichnet. Ein niedriger Maximum Drawdown zeigt, dass das Modell auch in Phasen negativer Marktentwicklungen vergleichsweise geringe Verluste verzeichnet, was einen wesentlichen Indikator für die Robustheit und Stabilität von Handelsstrategien darstellt (Vo & Ślepaczuk, 2022).

Die Kombination dieser Kennzahlen ermöglicht eine ganzheitliche Beurteilung der Modelle. Die annualisierte Rendite bewertet die durchschnittliche Performance, die Volatilität das Risiko, die Sharpe Ratio die Effizienz des Risiko-Ertrags-Verhältnisses und der Maximum Drawdown die Verlustanfälligkeit. Ein Modell gilt als ökonomisch tragfähig, wenn es

gleichzeitig eine überdurchschnittliche Rendite, eine moderate Volatilität, eine hohe Sharpe Ratio und einen niedrigen Maximum Drawdown aufweist.

### **Portfolio-Konstruktion und Analyse**

Neben der isolierten Bewertung einzelner Prognosemodelle ist es für die praktische Anwendung entscheidend, wie sich unterschiedliche Modelle kombinieren lassen, um ein robustes, diversifiziertes Portfolio aufzubauen. Die Portfolio-Konstruktion bildet somit den Übergang von der rein statistischen zur strategischen Evaluation der Prognosen.

### **Gleichgewichtetes Portfolio – die 1/N-Regel**

Ein einfacher, aber in der empirischen Finanzforschung vielfach validierter Ansatz ist die gleichgewichtete Portfolio-Konstruktion nach der sogenannten 1/N-Regel (DeMiguel, Garlappi & Uppal, 2009).

Dabei werden die zehn besten Prognosestrategien, gemessen an ihrer risikoadjustierten Performance (z. B. Sharpe Ratio oder Hit Ratio), mit identischen Gewichten kombiniert:

$$w_i = \frac{1}{N}, \quad \text{mit } N = 10$$

Dieses Verfahren hat den Vorteil, dass es modellunabhängig, robust und transparent ist. Da keine ex-ante Annahmen über die Korrelationen oder Risikostruktur getroffen werden, resultiert eine natürliche Diversifikation der Prognosen.

Der Diversifikationseffekt entsteht dadurch, dass sich modellindividuelle Prognosefehler teilweise gegenseitig ausgleichen, wodurch das Gesamtportfolio eine stabilere Performance aufweist.

### **Optimiertes Portfolio – Minimum-Variance-Ansatz**

Über die einfache Gleichgewichtung hinaus kann das Portfolio durch mathematische Optimierung gezielt verbessert werden.

Ein häufig genutzter Ansatz ist die Minimum-Variance-Optimierung nach Markowitz (1952), deren Ziel es ist, für eine gegebene Menge erwarteter Renditen das Risiko, gemessen durch die Varianz, zu minimieren:

$$w_{min} \quad w' \Sigma w \quad \text{unter den Nebenbedingungen} \quad \sum_{i=1}^N w_i = 1, \quad w_i \geq 0$$

Hierbei bezeichnet  $\Sigma$  die Kovarianzmatrix der Renditen der einzelnen Strategien und  $w$  den Vektor der Portfolioanteile.

Die Optimierung erfolgt in der Regel mittels quadratischer Programmierung, beispielsweise durch den Algorithmus *solve.QP()* aus dem *quadprog*-Paket (Turlach & Weingessel, 2019).

Die Constraints  $\sum w_i = 1$  und  $w_i \geq 0$  gewährleisten, dass keine Leerverkäufe erlaubt sind und das Portfolio vollständig investiert bleibt. Das Ergebnis ist eine Gewichtsverteilung, die das Risiko für die gegebene Renditestruktur minimiert, ein sogenanntes Minimum-Variance-Portfolio.

Die resultierenden Gewichte zeigen, welche Strategien maßgeblich zur Gesamtperformance beitragen. Hohe Gewichte deuten darauf hin, dass eine Strategie sowohl stabile als auch wenig korrelierte Erträge liefert. Niedrige Gewichte signalisieren hingegen entweder hohe Volatilität oder eine starke Korrelation mit anderen Strategien.

## **2.5 Anwendung in Handelsstrategien**

Die praktische Relevanz ökonometrischer Prognosemodelle zeigt sich insbesondere in ihrer Anwendung auf reale Finanzentscheidungen. Eine der bedeutendsten Einsatzmöglichkeiten liegt in der Entwicklung und Bewertung von Handelsstrategien, die auf den prognostizierten Bewegungen von Finanzmarktvariablen, wie den Renditen der 10-jährigen US-Treasuries, basieren. Ziel ist es, mithilfe statistischer Modelle profitable, risikoadjustierte Entscheidungen zu treffen und die Effizienz der Prognosen in einem ökonomischen Kontext zu überprüfen.

### **Konzept modellbasierter Handelsstrategien**

Eine modellbasierte Handelsstrategie nutzt Prognosen, um Kauf- oder Verkaufsentscheidungen für bestimmte Vermögenswerte abzuleiten. Die Grundidee besteht darin, systematische Signale aus der Modellvorhersage zu generieren, die über dem Zufall liegende Erträge ermöglichen.

Im Kontext der Anleihemärkte kann ein Prognosemodell dazu genutzt werden, zukünftige Veränderungen der Renditen abzuschätzen. Da Renditen und Anleihekurse in einem inversen Verhältnis zueinander stehen, ergibt sich daraus eine klare Handelslogik. Sinkt die prognostizierte Rendite, weist dies auf steigende Anleihepreise hin, wodurch eine Long-Position ökonomisch sinnvoll ist. Steigt hingegen die erwartete Rendite, ist mit fallenden Anleihekursen zu rechnen, sodass eine Short-Position die angemessenere Strategie darstellt.

Zur Bewertung der Wirtschaftlichkeit einer Prognosestrategie wird eine Backtesting-Analyse durchgeführt, bei der die Strategie rückwirkend auf historischen Daten getestet wird, um ihre Robustheit und praktische Umsetzbarkeit zu überprüfen. Dabei dienen die zuvor erläuterten Kennzahlen als zentrale Evaluationsgrößen.

## **2.6 Zielsetzung und Forschungsfragen**

Wie bereits in der Einleitung angedeutet, bewegt sich diese Arbeit im Spannungsfeld zwischen theoretischer Markteffizienz und praktischer Vorhersagbarkeit. Aufbauend auf den zuvor dargestellten theoretischen und methodischen Grundlagen soll nun untersucht werden, ob sich aus ökonometrischen Modellen tatsächlich prognostisch relevante Informationen über die Entwicklung langfristiger Zinsen ableiten lassen.

Das Ziel dieser Arbeit besteht darin, die Prognosefähigkeit unterschiedlicher ökonometrischer Modellklassen für die 10-jährige US-Staatsanleihenrendite systematisch zu bewerten und deren praktische Relevanz für Handels- und Portfolioentscheidungen zu

untersuchen. Die Analyse erfolgt auf Basis standardisierter Prognosegütemaße und Performancekennzahlen, um sowohl statistische als auch ökonomische Vorhersagequalität vergleichend zu beurteilen.

Dabei wird zwischen (1) univariaten Zeitreihenmodellen (ARIMA), (2) multivariaten fundamentalen Modellen und (3) hybriden Ansätzen (ARIMAX, ECM) unterschieden. Ziel ist es, zu bestimmen, welche Modellklasse bei gegebener Datenfrequenz und Prognosehorizont die höchste out-of-sample Prognosegüte sowie die beste risikoadjustierte Handelsperformance erzielt.

Aus der Zielsetzung dieser Arbeit ergeben sich drei zentrale Forschungsfragen, die den analytischen Rahmen der Untersuchung bilden. Erstens wird geprüft, welche Klasse ökonometrischer Modelle, univariate Zeitreihenmodelle, multivariate fundamentale Ansätze oder hybride Spezifikationen, die höchste Prognosegüte für die 10-jährige US-Staatsanleihenrendite erzielt. Damit wird untersucht, ob der Einbezug makroökonomischer Variablen wie Inflation, Arbeitsmarktindikatoren oder Term Spreads einen signifikanten Informationsgewinn gegenüber rein autoregressiven Modellen bietet. Die Bewertung der Prognosegüte erfolgt anhand standardisierter Fehlermaße wie dem Mean Absolute Error (MAE), der Hit Ratio (HR) und der Weighted Hitratio (WHR).

Zweitens wird analysiert, in welchem Umfang sich die aus den Modellen abgeleiteten Prognosen in einfache Handelsstrategien überführen lassen, die eine überlegene risikoadjustierte Rendite gegenüber einem passiven Buy-and-Hold-Ansatz erzielen. Die ökonomische Relevanz der Modelle wird dabei auf Basis von Performancemaßen wie der Sharpe Ratio, der Calmar Ratio, der annualisierten Rendite und dem maximalen Drawdown bewertet, um die praktische Umsetzbarkeit der Modellprognosen zu überprüfen.

Drittens wird untersucht, ob die Kombination mehrerer Prognosemodelle in einem diversifizierten Portfolio zu einer weiteren Verbesserung der Stabilität und Performance führt. Ziel ist es zu prüfen, ob sich durch die Kombination unterschiedlicher Modellklassen ein effizienteres und robusteres Ergebnis erzielen lässt als durch die isolierte Anwendung einzelner Prognoseansätze.

Zur Beantwortung dieser Forschungsfragen erfolgt die Implementierung einer Vielzahl ökonometrischer Modelle in der Programmiersprache R, deren Prognoseleistungen anschließend anhand etablierter statistischer Gütemaße systematisch analysiert und miteinander verglichen werden.

## **2.7 Stand der Forschung**

Die im vorherigen Abschnitt formulierten Forschungsfragen zielen darauf ab, die Prognosefähigkeit verschiedener Zeitreihenmodelle systematisch zu vergleichen und zu prüfen, in welchem Umfang makroökonomische Faktoren zur Verbesserung der Zinsprognose beitragen können. Um diese Fragen theoretisch zu fundieren und methodisch zu verankern, schließt sich nun eine vertiefte Betrachtung des aktuellen Forschungsstandes

an. Der folgende Abschnitt fasst zentrale wissenschaftliche Beiträge zusammen, die die Entwicklung und Anwendung ökonometrischer Prognosemodelle für Anleiherenditen maßgeblich geprägt haben. Die Forschung zur Prognose langfristiger Anleiherenditen stützt sich auf eine Reihe theoretisch und empirisch wegweisender Arbeiten.

Den Ausgangspunkt bildet die Studie von Nelson und Siegel (1987), die mit ihrem Drei-Faktoren-Modell, bestehend aus den Komponenten Level, Slope und Curvature, ein parsimonisches Framework zur Beschreibung der gesamten Zinsstruktur entwickelten. Dieses Modell bildet die theoretische Grundlage zahlreicher nachfolgender Ansätze und erklärt, wie sich die Dynamik der Renditekurve über verschiedene Laufzeiten hinweg systematisch erfassen lässt.

Darauf aufbauend erweiterten Diebold und Li (2006) den Ansatz, indem sie die Faktorisierung von Nelson und Siegel mit zeitreihenökonomischen Prozessen wie AR(1)- oder VAR(1)-Strukturen kombinierten. Ihr Dynamic Nelson-Siegel-Modell (DNS) etablierte sich als empirischer Benchmark, da es zeigte, dass einfache autoregressive Strukturen, ähnlich den im ARIMA-Ansatz verwendeten, die Bewegungen der Nelson-Siegel-Faktoren effektiv abbilden können.

Einen zentralen theoretischen Beitrag leisteten Engle und Granger (1987), die mit ihrem Konzept der Kointegration und dem darauf basierenden Error-Correction-Model (ECM) aufzeigten, dass nicht-stationäre Variablen, wie Zinsen und makroökonomische Indikatoren, dennoch eine stabile langfristige Beziehung aufweisen können. Das ECM wurde damit zu einem wichtigen Instrument, um kurzfristige Anpassungen an langfristige Gleichgewichte zu modellieren, ein Ansatz, der auch in dieser Arbeit aufgegriffen wird.

Empirisch bestätigten Castellani und dos Santos (o.J) die Überlegenheit solcher Gleichgewichtsmodelle. In ihrem Vergleich von ARIMA-, ECM- und KI-basierten Ansätzen zur Prognose der 10-jährigen US-Staatsanleihenrendite erzielte das ECM die geringsten Prognosefehler und übertraf sowohl klassische Zeitreihenmodelle, als auch künstliche Intelligenzverfahren.

Eine ähnliche Richtung verfolgten Bandholz, Clostermann und Seitz (2007), die das sogenannte Bond Yield Conundrum der 2000er-Jahre mithilfe eines ECM mit makroökonomischen Variablen untersuchten. Ihre Ergebnisse zeigten eine stabile Kointegrationsbeziehung zwischen Zinsen, Inflation und geldpolitischen Faktoren und belegten, dass ECM-Modelle Random-Walk-Prozesse signifikant übertreffen, ein zentrales Argument für den Einsatz makroökonomischer Kovariaten auch in ARIMAX-Modellen.

Einen stärker datengetriebenen Ansatz wählte Mönch (2005), der mit einem No-Arbitrage FAVAR-Modell (Factor-Augmented VAR) latente Zinsfaktoren mit aus über 160 US-Zeitreihen extrahierten Makrofaktoren kombinierte. Das Modell erzielte geringere Prognosefehler als das klassische Diebold-Li-Modell und belegte den Mehrwert makroökonomischer Variablen für die Vorhersage der Zinsstruktur.

In ähnlicher Weise integrierten Diebold, Rudebusch und Aruoba (2006) die Nelson-Siegel-Faktoren in ein dynamisches Zustandsraummodell mit makroökonomischen Größen wie Inflation, Kapazitätsauslastung und dem Federal Funds Rate. Ihre Ergebnisse dokumentierten signifikante Wechselwirkungen zwischen Zinsfaktoren und Makrovariablen und lieferten damit eine theoretische Grundlage für die Einbindung exogener Variablen in ARIMAX- und ECM-Modelle.

Bereits Campbell und Shiller (1989) hatten gezeigt, dass einfache Spread-Regressionen die Veränderung kurzfristiger Zinsen nur unzureichend erklären. Ihre empirischen Ergebnisse, negative Steigungskoeffizienten in Spread-Regressionen, verdeutlichten die begrenzte Prognosekraft linearer Ansätze und begründeten die Notwendigkeit komplexerer Modelle zur Erfassung der Zinsdynamik.

Aktuellere Arbeiten, etwa von Guidolin und Ionta (2025), untersuchten die Prognoseleistung von AR(1)- und VAR(1)-Modellen auf Basis von Nelson-Siegel-Faktoren und bewerteten diese mithilfe statistischer Vergleichstests wie dem Diebold-Mariano-Test und dem Model Confidence Set. Die Autoren kamen zu dem Ergebnis, dass multivariate Modelle, die die Dynamik mehrerer Faktoren simultan erfassen, die höchste Prognosegenauigkeit erreichen, ein Befund, der die Relevanz der multivariaten Modellierung auch für die 10-jährige US-Rendite unterstreicht.

Schließlich demonstrierte Zhang (2020) die praktische Anwendbarkeit klassischer Zeitreihenmodelle, indem er ein ARIMA-Modell für die Rendite 10-jähriger US-Staatsanleihen entwickelte und dieses erfolgreich auf Futures-Handelsstrategien übertrug. Das Ergebnis belegt den Nutzen einfacher, transparent parametrierter ARIMA-Modelle für praxisorientierte Prognosen und bildet zugleich einen Anknüpfungspunkt für die in dieser Arbeit durchgeführte Backtesting-Analyse.

### **Zentrale Erkenntnisse und Implikationen für die vorliegende Arbeit**

Aus den analysierten Studien lassen sich mehrere zentrale Erkenntnisse ableiten, die unmittelbar die methodische Ausrichtung dieser Arbeit begründen.

Erstens zeigen Nelson & Siegel (1987) sowie Diebold & Li (2006), dass die Dynamik der Zinsstruktur auch mit parsimonischen Modellen zuverlässig abgebildet werden kann, sofern die dominanten Faktoren (Level, Slope, Curvature) adäquat modelliert werden. Dieses Erkenntnis stützt den Einsatz autoregressiver Verfahren wie ARIMA als baseline-fähige Prognoseinstrumente.

Zweitens belegen Engle & Granger (1987) und Bandholz et al. (2007), dass langfristige Gleichgewichtsbeziehungen zwischen Zinsen und makroökonomischen Variablen bestehen und dass ECM-Ansätze diese stabil und empirisch nachvollziehbar erfassen. Damit rechtfertigt sich die Integration eines Fehlerkorrekturmodells als zentraler Bestandteil dieser Arbeit.

Drittens unterstreichen Mönch (2005) sowie Diebold, Rudebusch & Aruoba (2006), dass die Kombination latenter Zinsfaktoren mit makroökonomischen Indikatoren, analog zu einem ARIMAX-Ansatz, die Prognosequalität signifikant verbessert, da sie strukturelle Information jenseits rein statistischer Muster nutzt.

Und viertens verdeutlichen Campbell & Shiller (1989) sowie Castellani & dos Santos (o.J.), dass einfache Spread-Regressionen und KI-Modelle ohne ökonomische Fundierung nur begrenzt tragfähig sind, während ökonometrische Modelle mit klarer theoretischer Grundlage eine stabilere Prognoseleistung zeigen.

Schließlich zeigen Guidolin & Ionta (2025) sowie Zhang (2020), dass solche Modelle nicht nur akademisch valide, sondern auch in der praktischen Anwendung, etwa bei der Simulation regelbasierter Handelsstrategien, leistungsfähig sind.

Die Literatur bestätigt somit, dass die in dieser Arbeit verwendete Modelltrias aus ARIMA-, ARIMAX- und ECM-Ansätzen sowohl theoretisch fundiert als auch empirisch belastbar ist. Diese Kombination ermöglicht die Verbindung von statistischer Effizienz, ökonomischer Interpretierbarkeit und praktischer Anwendbarkeit, zentrale Eigenschaften im Hinblick auf die Out-of-Sample-Evaluierung und die Performanzmessung auf Basis realer Marktdaten. Im Unterschied zu früheren Studien, die sich überwiegend auf Daten bis zur Finanzkrise 2008 stützten, analysiert die vorliegende Arbeit die Prognosegüte über den Zeitraum 2006 bis 2025, der sowohl die Niedrigzinsphase und die Zinswende ab 2022 als auch die COVID-19-Pandemie umfasst. Darüber hinaus wird erstmals die ökonomische Verwertbarkeit der Modelle durch ein systematisches Backtesting von Handelsstrategien untersucht.

### 3 Datenbasis und empirisches Vorgehen

Nachdem im vorherigen Kapitel die theoretischen Grundlagen, Modellklassen und Gütemaße vorgestellt wurden, wendet sich dieses Kapitel der empirischen Umsetzung zu. Ziel ist es, die verwendeten Datenquellen, die Auswahl und Aufbereitung der Variablen sowie das methodische Vorgehen bei der Schätzung und Prognose der Modelle transparent und nachvollziehbar darzulegen.

Damit bildet das Kapitel die Schnittstelle zwischen der theoretischen Fundierung und der empirischen Analyse, es operationalisiert die in Kapitel 2 beschriebenen Konzepte für die konkrete Untersuchung der 10-jährigen US-Treasury-Rendite.

Zentrale Aufgabe ist die Sicherstellung einer konsistenten und ökonomisch fundierten Datengrundlage, die es erlaubt, sowohl kurzfristige als auch langfristige Prognosebeziehungen abzubilden. Entsprechend werden zunächst die verwendeten Datenquellen und Variablen beschrieben, anschließend die Transformations- und Testverfahren zur Sicherung der Stationarität erläutert und schließlich die Implementierung der verschiedenen Modellklassen dargestellt.

Die empirische Analyse basiert auf monatlichen Beobachtungen der 10-jährigen US-Treasury-Rendite als Zielvariable, ergänzt um eine Reihe makroökonomischer Einflussgrößen. Diese Variablen sind theoretisch in der Erwartungstheorie der Zinsstruktur (Diebold & Rudebusch, 2013, ch. 1; Guidolin & Ionta, 2025) und dienen dazu, die Informationsdynamik des US-Zinsmarkts umfassend zu erfassen.

Die Auswahl dieser Variablen erfolgt nach zwei Prinzipien:

- (1) Theoretische Relevanz, d. h. die Einbindung solcher Größen, die in der Literatur nachweislich mit langfristigen Zinssätzen in Beziehung stehen, und
- (2) Datenverfügbarkeit und Frequenzkonsistenz, um eine gleichmäßige Beobachtungsbasis über den gesamten Zeitraum zu gewährleisten.

#### 3.1 Datenakquise und -beschreibung

Die empirische Analyse basiert auf einem umfassenden Datensatz makroökonomischer Zeitreihen, der den Zeitraum vom 1. Oktober 1976 bis zum 30. September 2025 abdeckt. Diese lange Zeitspanne umfasst mehrere geldpolitische Regime, vom Volcker-Schock der frühen 1980er-Jahre über die Niedrigzinsphase nach der Finanzkrise 2008 bis hin zum jüngsten Hochzinsumfeld ab 2022 und ermöglicht damit eine robuste Untersuchung struktureller Veränderungen im US-Zinsmarkt (Diebold & Rudebusch, 2013, ch. 1; Mönch, 2006).

Die Auswahl der Variablen orientiert sich am aktuellen Stand der Forschung und umfasst Indikatoren, die in der Literatur als relevante Treiber langfristiger Anleiherenditen identifiziert wurden. Alle makroökonomischen Daten wurden der Datenbank Federal Reserve Economic Data (FRED, 2025) der Federal Reserve Bank of St. Louis entnommen. Der

Aktienmarktindex S&P 500 wurde über Yahoo Finance (2025) bezogen, während die für die Handelsstrategie verwendeten Daten des US Treasury Futures von Investing.com (2025) stammen. Er dient als Referenzinstrument zur Umsetzung der modellbasierten Signale in der Backtesting-Analyse.<sup>1</sup>

### **DGS10** – 10-jährige US-Staatsanleihenrendite

Die Rendite der 10-jährigen Treasury Note (DGS10) bildet die abhängige Variable aller Modelle. Sie gilt als zentrale Benchmark des globalen Finanzsystems und spiegelt die Erwartungen über Inflation, Wachstum und geldpolitische Ausrichtung wider (Guidolin & Ionta, 2025). Aufgrund ihres hohen Marktvolumens und ihrer Funktion als risikofreier Referenzzins wird sie häufig als Indikator für die langfristige Zinsstruktur herangezogen.

### **DGS2** – 2-jährige US-Staatsanleihenrendite

Die Rendite 2-jähriger US-Staatsanleihen (DGS2) repräsentiert den kurzfristigen Zins und spiegelt die unmittelbaren Markterwartungen bezüglich der geldpolitischen Ausrichtung der Federal Reserve wider. Aufgrund ihrer Sensitivität gegenüber geldpolitischen Entscheidungen wird die DGS2 in dieser Arbeit als kurzfristige Zinsvariable verwendet, um gemeinsam mit der langfristigen Rendite (DGS10) die Existenz einer langfristigen Gleichgewichtsbeziehung im Rahmen des Error-Correction-Models (ECM) nach dem Engle-Granger-Verfahren zu untersuchen. Dadurch ist sie ein zentraler Bestandteil der Cointegrationsanalyse zwischen kurz- und langfristigen Zinsen.

### **FEDFUNDS** – Federal Funds Rate

Die Federal Funds Effective Rate (FEDFUNDS) ist der kurzfristige Leitzins der US-Notenbank. Sie steuert das Zinsniveau im Interbankenmarkt und beeinflusst unmittelbar die kurzfristigen Erwartungen über zukünftige Zinsniveaus. Über die Erwartungstheorie der Zinsstruktur wirkt sie damit auf langfristige Renditen (Diebold, Rudebusch & Aruoba, 2006).

### **CPIAUCSL** – Consumer Price Index (CPI)

Der Consumer Price Index for All Urban Consumers (CPIAUCSL) misst die Preisentwicklung eines repräsentativen Warenkorbs und dient als Hauptindikator für die Inflationserwartungen. Da nominale Zinsen langfristig die erwartete Inflation widerspiegeln (Fisher-Gleichung), beeinflusst der CPI unmittelbar die langfristigen Renditen (Wang et al., 2025). In dieser Arbeit wird der jährliche prozentuale Veränderungswert (year-over-year) verwendet, um kurzfristige Schwankungen zu glätten und den zugrunde liegenden Inflationstrend adäquat abzubilden.

### **PPIACO** – Producer Price Index (PPI)

---

<sup>1</sup> Alle im Rahmen dieser Arbeit verwendeten Variablen wurden grafisch dargestellt; die entsprechenden Plots und Zeitreihencharts sind im Anhang A1 dargestellt.

Der Producer Price Index for All Commodities (PPIACO) erfasst Preisveränderungen auf der Produzentenebene. Er gilt als Frühindikator für künftige Verbraucherpreisinflation, da höhere Produktionskosten zeitverzögert auf Konsumentenpreise übergehen. Der PPI wird daher häufig in Modellen verwendet, die Inflationsdynamiken antizipieren (Ang & Piazzesi, 2003).

#### **INDPRO** – Industrial Production Index

Der Industrial Production Index (INDPRO) misst die reale Produktionsleistung des industriellen Sektors und gilt als Konjunkturindikator erster Ordnung. Ein steigender Wert signalisiert höhere Wirtschaftstätigkeit und tendenziell steigende Zinsen aufgrund zunehmender Kreditnachfrage (Ang & Piazzesi, 2003).

#### **UNRATE** – Arbeitslosenquote

Die Arbeitslosenquote (UNRATE) spiegelt den Zustand des Arbeitsmarktes und die Auslastung der Wirtschaft wider. Sinkende Arbeitslosigkeit geht häufig mit steigendem Lohndruck und Inflationserwartungen einher, was langfristige Renditen erhöht (Ang & Piazzesi, 2003).

#### **M2SL** – Geldmenge M2

Das Aggregat M2 (M2SL) umfasst Bargeld, Sichteinlagen und kurzfristige Spareinlagen. Eine Ausweitung der Geldmenge wirkt tendenziell inflationsfördernd und kann somit zu höheren langfristigen Zinsen führen. In geldpolitischen Analysen dient M2 als Indikator für Liquidität und expansive Maßnahmen (Hoesli et al. 2007).

#### **SP500** – Aktienmarktindex

Der Aktienindex S&P 500 (GSPC) repräsentiert die Marktkapitalisierung der 500 größten US-Unternehmen. Er spiegelt die Risikoneigung der Anleger und das Vertrauen in die wirtschaftliche Entwicklung wider. Ein steigender Aktienmarkt geht oft mit höheren Renditen risikofreier Anleihen einher, da Kapital von sicheren in risikoreichere Anlagen umgeschichtet wird (Castellani & dos Santos, o.J).

#### **Transformierte Variablen**

Zur Sicherstellung der Stationarität und Vergleichbarkeit werden mehrere Variablen in Wachstumsraten oder Differenzen transformiert:

**Tabelle 1:** Übersicht der transformierten Variablen und angewandten Transformationen

<b>Name</b>	<b>Variable</b>	<b>Art der Transformation</b>
10-jährige US-Staatsanleihenrendite	DGS10	erste Differenz
2-jährige US-Staatsanleihenrendite	DGS2	erste Differenz

Federal Funds Rate	FEDFUNDS	erste Differenz
Consumer Price Index (CPI)	CPIAUCSL	logarithmierte Differenzen (growth)
Producer Price Index (PPI)	PPIACO	logarithmierte Differenzen (growth)
Industrial Production Index	INDPRO	logarithmierte Differenzen (growth)
Arbeitslosenquote	UNRATE	erste Differenz
Geldmenge M2	M2SL	logarithmierte Differenzen (growth)
Aktienmarktindex	SP500	Monatliche Rendite (returns)

*Note.* Eigene Darstellung. Die Tabelle zeigt die für die Modellierung verwendeten makroökonomischen Variablen und deren jeweilige Transformationen zur Herstellung von Stationarität. Logarithmierte Differenzen entsprechen Wachstumsraten, erste Differenzen Veränderungen in Prozentpunkten.

### Wachstumsraten (log-Differenzen):

Zur Ermittlung der kontinuierlichen Wachstumsraten der Preis- und Aktivitätsvariablen werden logarithmierte Differenzen berechnet:

$$CPI_{growth} = 100 \times \Delta \ln(CPIAUCSL), PPI_{growth} = 100 \times \Delta \ln(PPIACO), \\ INDPRO_{growth} = 100 \times \Delta \ln(INDPRO), M2_{growth} = 100 \times \Delta \ln(M2SL).$$

Diese Transformationen messen die prozentualen Veränderungen und gelten in der ökonometrischen Literatur als Standardverfahren, um Verzerrungen durch nicht-stationäre Trends zu vermeiden (Box et al., 2016, ch. 2).

### Renditen

Die monatliche Rendite des Aktienmarkts (S&P 500) wird als logarithmierte Veränderung berechnet:

$$SP500_{returns} = 100 \times \ln\left(\frac{P_t}{P_{t-1}}\right).$$

Sie beschreibt die prozentuale Kursveränderung und dient als Indikator für das Marktsentiment sowie die Dynamik von Vermögenspreisen.

### Erste Differenzen

Für Zins- und Arbeitsmarktvariablen werden erste Differenzen gebildet, um kurzfristige Veränderungen zu isolieren und langfristige Trends zu eliminieren:

$$d_{DGS10} = \Delta DGS10, d_{DGS2} = \Delta DGS2, \\ d_{FEDFUNDS} = \Delta FEDFUNDS, d_{UNRATE} = \Delta UNRATE$$

Diese Transformation ist erforderlich, um stationäre Reihen zu erzeugen und die Voraussetzungen für die Schätzung von Modellen wie ARIMA, ARIMAX oder ECM zu erfüllen (Hyndman & Athanasopoulos, 2021, ch. 8.1).

### **3.2 Datenaufbereitung und Stationaritätstests**

Bevor die Modelle geschätzt und bewertet werden können, ist eine sorgfältige Aufbereitung der Daten erforderlich, um valide, reproduzierbare und verzerrungsfreie Ergebnisse zu gewährleisten. Ein zentraler Aspekt dabei ist die Vermeidung von Look-ahead-Bias, also der unbewussten Nutzung zukünftiger Informationen bei der Schätzung oder Prognose. Dieser Bias kann die Prognosegüte künstlich verbessern und ist daher insbesondere in ökonometrischen Zeitreihenanalysen streng zu vermeiden (Bailey et al., 2015).

Zur Sicherstellung einer methodisch sauberen Trennung wird der Datensatz in Trainings- und Testperiode unterteilt. Die Aufteilung der Stichprobe erfolgt chronologisch, um die zeitliche Abfolge der Daten beizubehalten. Dabei werden 60 Prozent des Gesamtzeitraums, von März 1977 bis März 2006, als Trainingsdaten genutzt, während die verbleibenden 40 Prozent, von April 2006 bis September 2025, als Testdaten dienen.

Diese Aufteilung stellt sicher, dass die Modelle auf einem langen, heterogenen Zeitraum geschätzt werden, der mehrere geldpolitische Regime umfasst, und anschließend auf einem rezenteren, unabhängigen Zeitraum getestet werden. Damit kann überprüft werden, ob die Modelle ihre Prognosegüte auch unter neuen Marktbedingungen beibehalten, etwa im Niedrigzinsumfeld nach 2008 oder in der Hochzinsphase ab 2022.

Die Entscheidung für ein 60/40-Split folgt der etablierten Praxis in der Finanz- und Makroökonomie. Sie bietet einen ausgewogenen Kompromiss zwischen Schätzpräzision und ausreichender Teststreuung (Clark & McCracken, 1999). Eine zu kurze Trainingsphase würde die Parameterschätzung instabil machen, während eine zu kurze Testphase die externe Validität der Ergebnisse einschränken würde. Durch diese Methodik wird gewährleistet, dass alle Prognosen out-of-sample, also außerhalb des Schätzzeitraums, erstellt werden.

#### **Master-Dataframe-Konzept**

Zur effizienten und transparenten Handhabung der zahlreichen Zeitreihenvariablen wird in der empirischen Umsetzung ein zentrales Master-Dataframe-Konzept verwendet. Dieses Konzept dient der einheitlichen Datenhaltung, Synchronisierung und Vorverarbeitung sämtlicher Variablen, um Inkonsistenzen in Zeitindizes, Frequenzen oder Datenlücken zu vermeiden.

Alle Variablen, Zielvariable, exogene Indikatoren und transformierte Reihen, werden zunächst in einem gemeinsamen Dataframe zusammengeführt. Dadurch entsteht eine konsistente Datenstruktur, die als zentrale Referenz für alle nachfolgenden Modellierungs- und Evaluationsschritte dient. Diese Vorgehensweise ermöglicht es, sämtliche

Berechnungen (Differenzen, log-Transformationen, Lag-Erzeugung, Normalisierung) innerhalb einer kohärenten Datenumgebung durchzuführen, was insbesondere für Reproduzierbarkeit und Nachvollziehbarkeit entscheidend ist. Es bildet somit das Rückgrat der empirischen Analyse. Es stellt sicher, dass sämtliche Modelle auf identischer Datengrundlage basieren, wodurch Vergleichbarkeit, Reproduzierbarkeit und methodische Sauberkeit gewährleistet werden.

Zur Erfassung zeitverzögerter Zusammenhänge wurden für alle relevanten Variablen Lag-Variablen der ersten bis dritten Periode erstellt. Auf diese Weise können sowohl kurzfristige Persistenzen als auch mittelfristige Transmissionseffekte, etwa verzögerte Reaktionen auf geldpolitische Impulse, erfasst werden (Enders, 2015, ch. 5; Diebold & Rudebusch, 2013, ch. 1). Die Generierung der Lags erfolgt automatisiert über die Funktion `lag.xts()` aus dem R-Paket `xts` (Ryan & Ulrich, 2008), wodurch der Zeitindex konsistent beibehalten wird.

Da viele makroökonomische Reihen in ihrem ursprünglichen Niveau nicht stationär waren, wurden sie durch erste Differenzen oder logarithmierte Wachstumsraten transformiert. Die Überprüfung der Stationarität erfolgte mithilfe des Augmented-Dickey-Fuller-Tests (ADF). Für alle im Modell verwendeten Reihen konnte die Nullhypothese einer Einheitswurzel auf einem Signifikanzniveau von 5 % verworfen werden, womit die Stationarität der Zeitreihen bestätigt ist. Die Ergebnisse der ADF-Tests sind im Anhang A2 dokumentiert. Da die Originalreihen nicht stationär sind, sind auch ihre Lags nicht stationär. Daher wurden in allen Modellen differenzierte Lags ( $\Delta X_{t-1}$ ,  $\Delta X_{t-2}$ ,  $\Delta X_{t-3}$ ) anstelle von Niveau-Lags verwendet, um Stationarität sicherzustellen. Für sämtliche Prognosemodelle kamen ausschließlich transformierte Daten, also differenzierte, logarithmierte, Wachstums- oder Renditereihen, zur Anwendung.

Fehlende Beobachtungen, die infolge unterschiedlicher Publikationsfrequenzen der zugrunde liegenden Datenquellen auftraten, wurden methodisch kontrolliert behandelt. Vereinzelt Ausfälle wurden aus dem Datensatz ausgeschlossen, während systematische Lücken durch eine zeitliche Beschränkung auf die längste gemeinsame Beobachtungsperiode bereinigt wurden. Auf eine Imputation fehlender Werte wurde bewusst verzichtet, um die Einführung potenziell künstlicher Strukturen in die Zeitreihen zu vermeiden und die empirische Integrität der Daten zu wahren.

Alle Zeitreihen wurden in einheitliche Monatsfrequenz überführt und im `xts`-Format geführt, um eine konsistente Zeitindexierung sicherzustellen. Nach jedem Verarbeitungsschritt erfolgte eine Prüfung auf identische Zeitstempel, eindeutige Einträge und regelmäßige Monatsabstände. Erst nach erfolgreicher Validierung dieser Konsistenzprüfungen wurden die Daten für die Modellierung freigegeben.

Nach Abschluss der Datenaufbereitung steht damit eine konsistente und methodisch saubere Grundlage für die empirische Modellierung zur Verfügung. Im folgenden Abschnitt werden die verwendeten Modellklassen und die zugrunde liegende Prognosestrategie im Detail vorgestellt.

### 3.3 Modellspezifikation und Prognosestrategie

Die empirische Analyse dieser Arbeit folgt einem systematischen Vorgehen, das von einfachen univariaten Zeitreihenmodellen über fundamentale Regressionsansätze bis hin zu fehlerkorrigierten und hybriden Modellen reicht. Ziel ist es, die Prognosegüte schrittweise zu steigern und zu überprüfen, ob die Integration makroökonomischer Fundamentalvariablen sowie langfristiger Gleichgewichtsbeziehungen einen substantziellen Mehrwert gegenüber rein datengetriebenen Ansätzen liefert. Die nachfolgende Darstellung orientiert sich an der chronologischen Abfolge der Modellierung im empirischen Teil (Kapitel 4) und dient als methodische Grundlage für die Interpretation der Ergebnisse.

#### Naive Prognose als Benchmark

Als Referenzmodell wird zunächst eine naive Prognose (Random Walk) verwendet, die annimmt, dass die beste Vorhersage der zukünftigen Renditeänderung die aktuelle Rendite ist:

$$\Delta \hat{y}_{t+1} = \Delta y_t$$

Dieses Modell stellt die Minimalanforderung an jede Prognose dar: Ein Modell sollte zumindest besser abschneiden als die Annahme „morgen ist wie heute“ (Hyndman & Athanasopoulos, 2021, ch. 3.1). Die naive Prognose dient als Effizienzmarkt-Benchmark, da sie impliziert, dass vergangene Bewegungen keine prädiktive Kraft besitzen, eine Annahme, die mit der schwachen Form der Effizienzmarkthypothese konsistent ist (Spiwoks et al., 2008).

#### Univariate ARIMA-Modelle

Die automatische Modellselektion über die Funktion *auto.arima()* aus dem R-Paket *forecast* (Hyndman & Athanasopoulos, 2021, ch. 8.5) identifizierte ein MA(1)-Modell als optimale univariate Spezifikation für die monatlichen Renditeänderungen der DGS10-Zeitreihe, weshalb dieses Modell für die weitere Analyse herangezogen wurde. Das Akaike Information Criterion (AIC) von 330,52 war das Niedrigste unter allen getesteten ARIMA(*p, d, q*)-Kombinationen mit  $p, q \in \{0,1,2,3\}$  (siehe Anhang A3).

Das MA(1)-Modell hat die Form:

$$\Delta y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{(t-1)}$$

wobei  $\Delta y_t$  die erste Differenz der Rendite darstellt,  $\mu$  den konstanten Term,  $\varepsilon_t$  einen White-Noise-Prozess und  $\theta_1$  den MA(1)-Koeffizienten bezeichnet.

Der geschätzte negative Koeffizient

$$\theta_1 = 0.158, \quad p < 0.001$$

weist auf eine kurzfristige Persistenz hin: Positive Schocks in Periode  $t - 1$  führen tendenziell zu positiven Anpassungen in Periode  $t$  (siehe Anhang A4).

Die Wahl des MA(1)-Modells gegenüber komplexeren Spezifikationen (z. B. ARMA(1,1) oder MA(2)) folgt dem Parsimonie-Prinzip. Komplexere Modelle verbesserten die Anpassungsgüte nur marginal, erhöhten jedoch das Risiko der Überanpassung (Box et al., 2016, ch. 4).

Zur Schätzung des MA(1)-Modells wurden zwei alternative Strategien verwendet, um unterschiedliche Anpassungsmechanismen an zeitliche Veränderungen zu berücksichtigen. Bei der rekursiven Prognose wird das Modell zunächst auf den Trainingsdaten von März 1977 bis März 2006 geschätzt und anschließend bei jeder neuen Beobachtung in der Testperiode mit allen bis dahin verfügbaren Daten neu kalibriert. Auf diese Weise wächst die Stichprobe kontinuierlich an, wodurch langfristige Muster und stabile Parameterstrukturen besser erfasst werden können (Neusser, 2025, ch. 14.6).

Im Gegensatz dazu basiert die Rolling-Window-Prognose auf einem festen Zeitfenster mit einer Länge von 2, 5, 10, 15 oder 20 Jahren, das sich fortlaufend über die Zeit verschiebt. Diese Methode reagiert flexibler auf Strukturbrüche und Regimewechsel und ermöglicht eine genauere Abbildung zeitvarianter Zusammenhänge (Neusser, 2025, ch. 14.6). Die Wahl der Rolling-Window-Längen von orientiert sich an der einschlägigen Prognoseliteratur, in der Zeitfenster zwischen 60 und 240 Monaten zur Analyse zeitvariabler Parameterbeziehungen üblich sind (Clark & McCracken, 1999). Kürzere Fenster ermöglichen eine stärkere Anpassung an kurzfristige Regimewechsel, während längere Fenster robustere Schätzungen über mehrere geldpolitische Phasen hinweg erlauben.

Für jede Strategie wurde eine 1-Schritt-voraus-Prognose ( $h=1$  Monat) erstellt, wobei das Modell bei jeder neuen Beobachtung neu geschätzt wurde, um adaptive Parameteranpassungen zu ermöglichen.

Neben dem festen MA(1)-Modell wurde auch eine adaptive ARMA-Modellierung untersucht, um die Prognosegüte durch eine dynamische Ordnungsanpassung zu verbessern. In der ersten Variante, der rekursiv optimierten ARMA-Schätzung, wird bei jeder neuen Prognose die optimale Modellordnung ARMA( $p,q$ ) mithilfe eines manuell programmierten Verfahrens neu bestimmt. Dadurch kann das Modell auf zeitvariante Dynamiken reagieren und strukturelle Veränderungen im Verlauf der Stichprobe adaptiv erfassen. In der zweiten Variante, dem ARMA-Rolling-Window-Ansatz, erfolgt die Schätzung innerhalb eines festen, über die Zeit gleitenden Fensters von 20 Jahren. Für jeden Zeitpunkt wird dabei erneut die optimale Modellordnung ( $p,q$ ) bestimmt, um eine bestmögliche Anpassung an lokale Muster und kurzfristige Regimewechsel zu gewährleisten.

**Tabelle 2:** Übersicht der univariaten Prognosemodelle (ARIMA-Klasse)

Modelle	Eigenschaften	Variablen
Naive	„morgen ist wie heute“	DSG10; Lag-1
ARIMA		
MA(1) Rekursiv	Rekursive Prognose	DSG10; Lag-1
MA(1) Rol. 2J	Rolling-Window über 2 Jahre	DSG10; Lag-1
MA(1) Rol. 5J	Rolling-Window über 5 Jahre	DSG10; Lag-1
MA(1) Rol. 10J	Rolling-Window über 10 Jahre	DSG10; Lag-1
MA(1) Rol. 15J	Rolling-Window über 15 Jahre	DSG10; Lag-1
MA(1) Rol. 20J	Rolling-Window über 20 Jahre	DSG10; Lag-1
ARMA Rol. Fest 20J	Rolling-Window über 20 Jahre, manuelle feste p,q Wahl	DSG10; Lag-1
ARMA Rek. Opt.	Rekursive Prognose, immer neue p,q Wahl	DSG10; Lag-1
ARMA Rol. Opt. 20J	Rolling-Window über 20 Jahre, immer neue p,q Wahl	DSG10; Lag-1

*Note.* Eigene Darstellung. Die Tabelle zeigt die in der Arbeit implementierten univariaten Prognosemodelle für die Renditeänderungen der 10-jährigen US-Staatsanleihe (DGS10). Unterschieden wird zwischen rekursiven und rollierenden Schätzansätzen mit unterschiedlichen Fensterlängen. MA(1) bezeichnet das Moving-Average-Modell erster Ordnung.

## Fundamentale Regressionsmodelle

Da ARIMA-Modelle lediglich vergangene Renditebewegungen abbilden und keine makroökonomischen Informationen enthalten, wurde im nächsten Schritt geprüft, ob sich die Prognosegüte durch die Einbeziehung fundamentaler Einflussgrößen verbessern lässt. Zu diesem Zweck wurden fünf multivariate Regressionsmodelle mit unterschiedlichen Kombinationen exogener Variablen spezifiziert.

Die allgemeine Form lautet:

$$\Delta y_t = \beta_0 + \sum_{i=1}^k \beta_i X_{i,t-1} + \varepsilon_t$$

wobei  $X_{i,t-1}$  die um eine Periode verzögerten exogenen Variablen darstellen. Die Verwendung von Lags gewährleistet, dass ausschließlich zum Prognosezeitpunkt  $t$  verfügbare Informationen einfließen (Carnot et al., 2005, ch. 4.2).

**Tabelle 3:** Übersicht der fundamentalen Regressionsmodelle (multivariat)

Modelle	Eigenschaften	Variablen
Fundamental		
Modell_1	Alle verfügbaren Variablen; Lag-1	DSG10, DGS2, FEDFUNDS, CPIAUCSL, PPIACO, INDPRO, UNRATE, M2SL und SP500
Modell_2	fundamentale Treiber kurzfristiger Zinsbewegungen; Lag-1	DGS2, FEDFUNDS, M2SL, PPIACO und SP500
Modell_3	Alle verfügbaren Variablen; Lag-1 bis Lag-3	DSG10, DGS2, FEDFUNDS, CPIAUCSL, PPIACO, INDPRO, UNRATE, M2SL und SP500
Modell_4	Nur S&P 500, aufgrund hoher statistischer Signifikanz; Lag-1	SP500
Stepwise	Auswahl auf Basis automatischer Variablenselektion durch stepAIC() Funktion (explorativ); Lag-1 bis Lag-3	DSG10, DGS2, FEDFUNDS, CPIAUCSL, PPIACO, INDPRO, UNRATE, M2SL und SP500

Note. Eigene Darstellung. Die Tabelle zeigt die in der Arbeit spezifizierten fundamentalen Modelle zur Prognose der Renditeänderungen der 10-jährigen US-Staatsanleihe (DGS10). Je nach Modellvariante wurden unterschiedliche Kombinationen makroökonomischer Variablen (DGS10, DGS2, FEDFUNDS, M2SL, CPIAUCSL, PPIACO, INDPRO, UNRATE, SP500) sowie verschiedene Lag-Strukturen berücksichtigt. Stepwise bezeichnet ein automatisches Modellselektionsverfahren auf Basis der Funktion stepAIC() aus dem R-Paket MASS (Venables & Ripley, 2002), das mithilfe des Akaike Information Criterion (AIC) die prädiktiv relevanten Variablen identifiziert. Der Signifikanznachweis zu Modell\_4 befindet sich im Anhang A5.

### Fehlerkorrekturmodelle (Error Correction Models, ECM)

Die fundamentalen Regressionsmodelle erfassen zwar kurzfristige Zusammenhänge, ignorieren jedoch potenzielle langfristige Gleichgewichtsbeziehungen. Das Konzept der Kointegration (Engle & Granger, 1987) besagt, dass nicht-stationäre Zeitreihen eine gemeinsame stochastische Tendenz aufweisen können, was durch einen stationären Fehlerkorrekturterm (Error Correction Term, ECT) modelliert wird (Granger et al., 1998).

Der Engle-Granger-Test ergab eine signifikante Kointegration zwischen DGS10 und DGS2. Der Augmented-Dickey-Fuller-Test (ADF) der geschätzten Residuen liefert eine Teststatistik von  $-3,478$  ( $p = 0,044$ ), womit die Nullhypothese einer Einheitswurzel verworfen werden kann (siehe Anhang A6). Dieses Ergebnis stützt die theoretische Erwartungshypothese der Zinsstruktur, nach der langfristige Zinsen durch erwartete kurzfristige Zinsen determiniert werden (Diebold & Rudebusch, 2013, ch. 1). Die ECM-Spezifikation lautet:

$$\Delta y_t = \alpha + \gamma \cdot ECT_{t-1} + \sum_{i=1}^k \beta_i \Delta X_{i,t-1} + \varepsilon_t$$

wobei

$$ECT_{t-1} = y_{t-1} - \hat{\theta}_0 - \hat{\theta}_1 \cdot DGS_{t-1}$$

die Abweichung vom langfristigen Gleichgewicht darstellt. Der Anpassungskoeffizient  $\gamma$  misst die Geschwindigkeit, mit der die Rendite wieder in Richtung des langfristigen Gleichgewichts (Fair Value) konvergiert.

**Tabelle 4:** Übersicht der Error-Correction-Modelle (ECM-Klasse)

Modelle	Eigenschaften	Variablen
ECM		
ECM_1	Nur Error Correction Term; Lag-1	ECT
ECM_2	Sequenzielle Variation der Variablen und Vergleich der resultierenden Prognosegütern: Lag-1	ECT, DGS10, DGS2 und SP500
ECM_3	Nur Error Correction Term und S&P 500; Lag-1	ECT und SP500
ECM_4	Auswahl auf Basis automatischer Variablenselektion durch stepAIC() Funktion (explorativ); Lag-1 bis Lag-3	ECT, DSG10, DGS2, FEDFUNDS, CPIAUCSL, PPIACO, INDPRO, UNRATE, M2SL und SP500
ECM_5	Alle verfügbaren Variablen; Lag-1 bis Lag-3	ECT, DSG10, DGS2, FEDFUNDS, CPIAUCSL, PPIACO, INDPRO, UNRATE, M2SL und SP500
Manual Final	Error Correction Term und Variablen aus Modell_2; Lag-1	ECT, DGS2, FEDFUNDS, M2SL, PPIACO und S&P 500

Note. Eigene Darstellung. Die Tabelle zeigt die in der Arbeit spezifizierten Error-Correction-Modelle (ECM) zur Prognose der Renditeänderungen der 10-jährigen US-Staatsanleihe (DGS10). In allen Modellen wird der Error-Correction-Term (ECT) als langfristige Kointegrationskomponente integriert. Die Modellvarianten unterscheiden sich hinsichtlich der einbezogenen fundamentalen Variablen (DGS10, DGS2, FEDFUNDS, M2SL, CPIAUCSL, PPIACO, INDPRO, UNRATE, SP500) sowie der gewählten Lag-Struktur. ECM\_4 verwendet eine automatische Variablenselektion auf Basis der Funktion stepAIC() aus dem R-Paket MASS (Venables & Ripley, 2002), während ECM\_2 empirisch über eine sequentielle Variation der Variablen bestimmt wurde. Das Manual Final-Modell kombiniert den Error-Correction-Term mit den signifikanten Variablen aus Modell 2 und dient als finale Vergleichsspezifikation.

### Hybride ARIMAX-Modelle

Die ARIMAX-Klasse kombiniert die autoregressive Struktur der ARIMA-Modelle mit exogenen Variablen (Shumway & Stoffer, 2018, ch. 5.6). Dies erlaubt die Integration fundamentaler Informationen bei gleichzeitiger Erfassung vergangener Renditebewegungen:

$$\Delta y_t = \mu + \theta_1 \varepsilon_{t-1} + \sum_{i=1}^k \beta_i X_{i,t-1} + \varepsilon_t$$

Dabei bezeichnet  $\Delta y_t$  die erste Differenz der Zielgröße, wodurch Stationarität hergestellt wird. Das konstante Glied  $\mu$  repräsentiert den Mittelwert der Serie, während  $\theta_1 \varepsilon_{t-1}$  die Moving-Average-Komponente erster Ordnung abbildet und kurzfristige Schocks der Vorperiode berücksichtigt. Der Summenausdruck  $\sum_{i=1}^k \beta_i X_{i,t-1}$  umfasst die verzögerten exogenen Variablen, die fundamentale Einflussfaktoren wie Zinsniveau, Inflation oder Produktionswachstum modellieren. Schließlich steht  $\varepsilon_t$  für den weißen Rauschterm, der unbeobachtete, zufällige Störungen der Zeitreihe beschreibt.

**Tabelle 5:** Übersicht der hybriden ARIMAX-Modelle (kombinierte Zeitreihen- und Fundamentaldatenansätze)

Modelle	Eigenschaften	Variablen
ARIMAX		
ARIMAX	Rekursives MA(1) mit die Variablen aus ECM_2; Lag-1	ECT, DGS10, DGS2 und SP500
ARIMAX Rol. 2J	Rolling-Window über 2 Jahre, MA(1) und die Variablen aus ECM_2; Lag-1	ECT, DGS10, DGS2 und SP500
ARIMAX Rol. 5J	Rolling-Window über 5 Jahre, MA(1) und die Variablen aus ECM_2; Lag-1	ECT, DGS10, DGS2 und SP500
ARIMAX Rol. 10J	Rolling-Window über 10 Jahre, MA(1) und die Variablen aus ECM_2; Lag-1	ECT, DGS10, DGS2 und SP500
ARIMAX Rol. 15J	Rolling-Window über 15 Jahre, MA(1) und die Variablen aus ECM_2; Lag-1	ECT, DGS10, DGS2 und SP500
ARIMAX Rol. 20J	Rolling-Window über 20 Jahre, MA(1) und die Variablen aus ECM_2; Lag-1	ECT, DGS10, DGS2 und SP500

Note. Eigene Darstellung. Die Tabelle zeigt die in der Arbeit spezifizierten ARIMAX-Modelle, welche die autoregressive Struktur der ARIMA-Klasse mit den fundamentalen Variablen aus ECM 2 (ECT, DGS10, DGS2 und SP500) kombinieren. Alle Modelle basieren auf einem MA(1)-Prozess und unterscheiden sich hinsichtlich des Schätzansatzes (rekursiv vs. rollierend) sowie der Länge des verwendeten Rolling-Windows (2 bis 20 Jahre).

Die ARIMAX-Modelle stellen eine Zwischenlösung dar. Sie sind komplexer als reine ARIMA-Ansätze, aber parsimonischer als vollständige ECM-Spezifikationen.

### Zusammenfassung der Modellhierarchie

Die 27 untersuchten Modellspezifikationen lassen sich wie folgt hierarchisch ordnen:

**Tabelle 6:** Übersicht der verwendeten Modellklassen und Spezifikationen

Modellklasse	Anzahl Varianten	Zentrale Eigenschaft
Naive Prognose	1	Effizienzmarkt-Benchmark
ARIMA/MA(1)	10	Rein datengetrieben, keine fundamentalen Variablen
Fundamentale Regression	5	Kurzfristige makroökonomische Zusammenhänge
ECM	6	Langfristige Gleichgewichtsbeziehungen + Fundamentaldaten
ARIMAX	6	Hybridansatz: Autoregression + Fundamentaldaten

*Note.* Eigene Darstellung. Die Tabelle fasst die in dieser Arbeit implementierten Modellklassen zur Prognose der Renditeänderungen 10-jähriger US-Staatsanleihen zusammen.

Die Modellarchitektur folgt einer schrittweisen Komplexitätssteigerung von parsimonen, rein datengetriebenen Ansätzen hin zu zunehmend strukturell fundierten Modellen. Jede Stufe bildet dabei eine spezifische Forschungshypothese ab und erweitert die Informationsbasis systematisch.

In einem ersten Schritt prüfen die univariaten ARIMA-Modelle, ob sich aus den vergangenen Renditebewegungen eigenständige Prognosemuster ableiten lassen. Damit wird die zentrale Frage adressiert, ob die Zeitreihe selbst Informationen über ihre zukünftige Entwicklung enthält. Darauf aufbauend untersuchen die fundamentalen Regressionsmodelle, ob makroökonomische Einflussgrößen zusätzliche Erklärungs- und Prognosekraft besitzen. Die Fehlerkorrekturmodelle erweitern diesen Ansatz um die Dimension langfristiger Gleichgewichtsbeziehungen. Sie prüfen, ob zwischen den Renditen und fundamentalen Variablen stabile Kointegrationszusammenhänge bestehen, die eine prädiktive Relevanz aufweisen. Abschließend kombinieren die hybriden ARIMAX-Modelle beide Perspektiven, indem sie sowohl die autoregressive Dynamik der Rendite als auch fundamentale Einflussfaktoren integrieren. Auf diese Weise wird untersucht, ob sich durch die Verbindung beider Informationsquellen eine höhere Prognosegüte erzielen lässt.

### **Prognosestrategie und Out-of-Sample-Evaluation**

Alle Modelle wurden konsistent mit einer 1-Schritt-voraus-Prognose ( $h=1$  Monat) evaluiert. Die Testperiode (April 2006 – September 2025) umfasst 234 Beobachtungen, was eine robuste Bewertung der Out-of-Sample-Performance ermöglicht (Clark & McCracken, 2011, ch. 2.10).

Für die Erstellung jeder Prognose  $\hat{y}_{t+1}$  wurde ein standardisiertes Verfahren angewendet, das den Ablauf einer rekursiven beziehungsweise rollierenden Schätzung abbildet. Zunächst erfolgt die Schätzung des Modells entweder auf Basis aller bis zum Zeitpunkt  $t$  verfügbaren Beobachtungen (rekursiv) oder innerhalb eines festen Zeitfensters (Rolling Window), um potenzielle Strukturveränderungen im Zeitverlauf zu berücksichtigen. Anschließend wird auf Grundlage der geschätzten Parameter die Ein-Schritt-Prognose für den folgenden Zeitraum  $\hat{y}_{t+1}$  generiert. Nach Ablauf des Prognoseintervalls wird der tatsächlich eingetretene Wert  $y_{t+1}$  beobachtet und mit der vorhergesagten Rendite verglichen, um die Vorhersagegüte zu evaluieren. Abschließend wird das Modell um den neu beobachteten Wert erweitert und erneut geschätzt, wodurch ein fortlaufender, adaptiver Lernprozess entsteht, der die kontinuierliche Aktualisierung der Prognosen ermöglicht.

## 4 Empirische Ergebnisse

Die folgenden Abschnitte dokumentieren die empirische Umsetzung der in Kapitel 2 theoretisch fundierten Prognosemodelle. Die Darstellung folgt dabei konsequent der chronologischen Vorgehensweise der Datenanalyse. Beginnend mit der Datenaufbereitung, über die Modellierung und Schätzung, bis hin zur Bewertung der resultierenden Strategien und Portfolioeffekte.

### 4.1 Univariate ARIMA-Modelle: Baseline-Performance

Zur Bewertung der Prognosequalität der univariaten Modelle werden im Folgenden zentrale Gütemaße herangezogen. Tabelle 7 fasst die Ergebnisse der Out-of-Sample-Prognosen zusammen und vergleicht die MA(1)-Modelle unterschiedlicher Schätzansätze mit der Naiven Benchmark-Prognose.

**Tabelle 7:** Prognosegüte der univariaten Modelle (ARIMA/MA-Klasse)

Modell	MAE	MSE	Cor	Hit Ratio	WHR
Naive Prognose	0.264	0.1177	0.0951	0.498	0.490
MA(1) Rekursiv	0.1936	0.0644	0.1024	0.4978	0.5036
MA(1) Rol. 2J	0.2033	0.0728	-0.0965	0.4410	0.4575
MA(1) Rol. 5J	0.1964	0.0670	-0.0630	0.4672	0.4889
MA(1) Rol. 10J	0.1944	0.0657	-0.0100	0.4891	0.5321
MA(1) Rol. 15J	0.1934	0.0647	0.0794	0.4891	0.5081
MA(1) Rol. 20J	0.1937	0.0647	0.0822	0.4978	0.5111
ARMA Rol. Fest 20J	0.1943	0.0644	0.1099	0.5197	0.5479
ARMA Rek. Opt.	0.1921	0.0631	0.1656	0.5590	0.5972
ARMA Rol. Opt. 20J	0.2053	0.0702	-0.0220	0.4541	0.4699

*Note.* Eigene Darstellung. Die Tabelle zeigt die Out-of-Sample-Prognosegüte der univariaten Modelle im Vergleich zur Naiven Prognose als Effizienzmarkt-Benchmark. Bewertet wurden die Modelle anhand des Mean Absolute Error (MAE), Mean Squared Error (MSE), der Korrelation zwischen vorhergesagten und tatsächlichen Renditeänderungen (Cor) sowie der Trefferquote (Hit Ratio) und der gewichteten Trefferquote (WHR). Geringere Fehlerwerte (MAE, MSE) und höhere Hit-Ratios weisen auf eine bessere Prognosequalität hin.

#### Naive Prognose als Effizienzmarkt-Benchmark

Die naive Prognose (Random Walk) erzielt einen MAE von 0,264 und eine Hit Ratio von 49,8 %, was nahe am Zufallsniveau liegt und mit der schwachen Form der Effizienzmarkthypothese konsistent ist. Die Weighted Hit Ratio (WHR) beträgt 49,0 %, was

darauf hindeutet, dass die naive Prognose auch bei starken Bewegungen keine systematische Richtungsgenauigkeit aufweist.

### **MA(1)-Modell: Erste signifikante Verbesserungen**

Das über die gesamte Testperiode hinweg rekursiv geschätzte MA(1)-Modell mit fixierter Ordnung zeigt eine klare Verbesserung gegenüber der naiven Benchmark. Der mittlere absolute Fehler (MAE) beträgt 0,1936 und liegt damit rund 26,6 % unter dem Wert der Random-Walk-Prognose. Auch der mittlere quadratische Fehler (MSE) verringert sich um 45,3 % auf 0,0644, was auf eine deutlich präzisere Anpassung des Modells hinweist. Die Richtungsgenauigkeit bleibt auf einem vergleichbaren Niveau: Die Hit Ratio erreicht 49,8 % und liegt damit nur geringfügig über der Naive-Prognose (49,8 %). Ein leichter Zugewinn zeigt sich hingegen bei der Weighted Hit Ratio, die mit 50,4 % einen um 1,4 Prozentpunkte höheren Wert erzielt und damit auf eine geringfügig verbesserte richtungsgewichtete Prognosequalität hinweist.

Der geschätzte positive MA(1)-Koeffizient ( $\theta_1 = 0.158$ ,  $p < 0.001$ ) weist auf eine kurzfristige Persistenz hin: Positive Schocks in Periode  $t - 1$  führen tendenziell zu positiven Anpassungen in Periode  $t$  (siehe Anhang A4).

### **Rolling-Window-Strategien: Adaptive Modellierung**

Die Verwendung rollierender Schätzfenster zeigt, dass längere Fensterbreiten tendenziell stabilere Ergebnisse liefern. Eine Fensterlänge von 15 bis 20 Jahren bietet die beste Balance zwischen Anpassungsfähigkeit und Schätzstabilität, während sehr kurze Zeiträume (z. B. 2 Jahre) aufgrund geringer Stichprobengröße zu schwankenden Parametern führen. Das leistungsfähigste Rolling-Window-MA(1)-Modell mit 15 Jahren erzielt einen MAE von 0,1934 und eine WHR von 50,81 %, während das 20-Jahres-Fenster mit einem MAE von 0,1937 und einer WHR von 51,11 % eine ähnlich gute Performance liefert. Damit reduzieren beide Modelle die Fehler gegenüber der naiven Prognose um etwa 26 % und zeigen eine leichte Steigerung der Richtungsgenauigkeit.

### **ARMA-Modell mit rollierendem Fenster und fester Ordnung**

Die AIC-basierte manuelle (keine Verwendung der `auto.arima()` Funktion) Ordnungsbestimmung ergab ein ARMA(1,1)-Modell als beste Kombination aus Anpassungsfähigkeit und Modellparsimony (AIC = 332,08). Diese wurde am Anfang bestimmt und dann für die gesamte Schätzung verwendet. Das ARMA(1,1)-Modell wurde mit einem 20-jährigen rollierenden Fenster geschätzt. Es erreicht einen MAE von 0,1943 (-26,4 % vs. naive Prognose), einen MSE von 0,0644 (-29,5 %) und eine Hit Ratio von 51,97 %. Die Weighted Hit Ratio liegt bei 54,79 % und damit deutlich über der Benchmark. Damit übertrifft das Modell die MA(1)-Spezifikation leicht und weist auf stabile kurzfristige Korrekturmuster hin, die in einem adaptiven Schätzfenster konsistent erfasst werden.

### **ARMA-Modell mit rekursiv anwachsendem Fenster und optimaler Ordnung**

In der rekursiv anwachsend geschätzten Variante wurde die Modellordnung (p,q) zu jedem Zeitpunkt neu über das AIC-Kriterium bestimmt. Der MAE liegt bei 0,1921, der MSE bei 0,0631, die Korrelation zwischen Prognose und Realisation bei 0,1656, die Hit Ratio bei 55,9 % und die Weighted Hit Ratio bei 59,72 %. Diese Werte verdeutlichen, dass eine dynamische Anpassung der Modellordnung an den verfügbaren Informationsstand die Erfassung kurzfristiger Zinsdynamiken optimiert. Besonders der Anstieg der Weighted Hit Ratio zeigt, dass das Modell in Phasen größerer Marktschwankungen robustere Richtungsprognosen liefert.

### **ARMA-Modell mit rollierendem Fenster und optimaler Ordnung**

Wird die Modellordnung zusätzlich innerhalb eines 20-jährigen rollierenden Fensters fortlaufend optimiert, nehmen die Effizienzgewinne ab. Das Modell erzielt einen MAE von 0,2053, einen MSE von 0,0702, eine Hit Ratio von 45,4 % und eine Weighted Hit Ratio von 47,0 %. Diese Ergebnisse zeigen, dass eine zu häufige Re-Kalibrierung des Modells zu Instabilitäten führen kann, insbesondere bei begrenzten Beobachtungsumfängen innerhalb der rollierenden Fenster. Während moderate Rolling-Strategien die Anpassungsfähigkeit erhöhen, führt eine zu starke Flexibilisierung nicht zwangsläufig zu besseren Vorhersagen.

## **4.2 Fundamentale Regressionsmodelle: Integration makroökonomischer Informationen**

In diesem Abschnitt werden fundamentale Regressionsmodelle vorgestellt, die zur Integration makroökonomischer Informationen in die Prognose der US-Treasury-Rendite dienen. Um den Einfluss zentraler makroökonomischer Determinanten auf die Entwicklung der US-Treasury-Rendite abzubilden, wurden im nächsten Schritt fundamentale Regressionsmodelle geschätzt. Diese Modelle erweitern die zuvor betrachteten univariaten Ansätze um exogene Einflussgrößen. Die in Kapitel 3.3 beschriebenen Spezifikationen werden im Folgenden hinsichtlich ihrer Prognosegüte miteinander verglichen. Die Tabelle 8 fasst die zentralen Ergebnisse zusammen.

**Tabelle 8:** Prognosegüte der fundamentalen Regressions- und ECM-Modelle

Modell	<i>MAE</i>	MSE	Cor	Hit Ratio	WHR
Modell_1	0,2115	0.0767	0.1768	0,524	0.576
Modell_2	0,1941	0.0636	0.2631	0,568	0.5739
Modell_3	0,2588	0.1475	0.0553	0,493	0,555
Modell_4	0,1908	0.0619	0.2123	0,550	0,562

---

Stepwise	0,2426	0.1247	0.0803	0.4803	0.5452
----------	--------	--------	--------	--------	--------

---

*Note.* Eigene Darstellung. Die Tabelle zeigt die Out-of-Sample-Prognoseergebnisse der multivariaten Modelle im Vergleich verschiedener Spezifikationen. Bewertet wurden die Modelle anhand des Mean Absolute Error (MAE), Mean Squared Error (MSE), der Korrelation zwischen prognostizierten und tatsächlichen Renditeänderungen (Cor) sowie der Trefferquote (Hit Ratio) und der gewichteten Trefferquote (WHR). Geringere Fehlerwerte (MAE, MSE) und höhere Hit Ratios deuten auf eine bessere Prognosequalität hin.

Die Ergebnisse zeigen, dass insbesondere Modell\_2 und Modell\_4 die beste Prognoseleistung aufweisen. Modell\_4 erzielt mit einem MAE von 0,1908 und einer Weighted Hit Ratio von 56,2 % die insgesamt höchste Effizienz und verdeutlicht, dass Aktienmarktentwicklungen (SP500) eine zentrale Rolle bei der kurzfristigen Vorhersage von Renditebewegungen spielen. Auch Modell\_2, das zusätzlich geldpolitische und monetäre Einflussgrößen wie FEDFUNDS und M2SL integriert, erreicht eine hohe Prognosegüte (MAE = 0,1941; WHR = 57,4 %) und bestätigt die Relevanz makroökonomischer Faktoren. Modell 1, das ausschließlich fundamentale Kernvariablen umfasst, zeigt mit einem MAE von 0,2115 und einer Weighted Hit Ratio von 57,6 % eine solide, aber geringfügig schwächere Performance. Die hohe Richtungsgenauigkeit deutet darauf hin, dass bereits ein kompaktes Set an Basisvariablen wesentliche Teile der Renditedynamik erfassen kann, während die zusätzlichen Faktoren in Modell 2 und 4 vor allem zur weiteren Feinanpassung beitragen. Demgegenüber verschlechtert Modell\_3 mit der Integration mehrfach gelagter Variablen (Lag 1 bis 3) die Vorhersagequalität deutlich (MAE = 0,2588), was auf Multikollinearität und eine Überparametrisierung hindeutet (Greene, 2002, ch. 11.4). Die automatische Stepwise-Regression liefert ebenfalls keine Leistungssteigerung (MAE = 0,2426; WHR = 54,5 %).

Die Schätzung von Modell\_2 zeigt, dass insbesondere kurzfristige Zinsen und realwirtschaftliche Indikatoren die Dynamik der zehnjährigen US-Treasury-Rendite bestimmen. Die zweijährige Rendite ( $d\_DGS2\_lag1$ ,  $\beta = 0,092$ ;  $p = 0,028$ ) wirkt signifikant positiv, was die Erwartungstheorie der Zinsstruktur bestätigt. Steigende kurzfristige Zinsen führen somit zu höheren langfristigen Renditen. Auch der Erzeugerpreisindex ( $PPI\_g\_lag1$ ,  $\beta = 0,066$ ;  $p = 0,028$ ) und die Aktienmarktentwicklung ( $SP500\_r\_lag1$ ,  $\beta = 0,015$ ;  $p = 0,003$ ) zeigen signifikant positive Effekte und verdeutlichen, dass sowohl Preis- als auch Wachstumserwartungen auf die Renditeentwicklung durchschlagen. Die detaillierten Schätzergebnisse, einschließlich der Regressionskoeffizienten, Signifikanzniveaus und Diagnostikstatistiken, sind im Anhang A5.2 dokumentiert. Dagegen bleiben geldpolitische Impulse ( $d\_FEDFUNDS\_lag1$ ,  $\beta = 0,008$ ;  $p = 0,816$ ) und die Geldmengenentwicklung ( $M2\_g\_lag1$ ,  $\beta = 0,079$ ;  $p = 0,189$ ) statistisch insignifikant, was auf eine nur begrenzte kurzfristige Transmission geldpolitischer Maßnahmen in die langfristigen Zinsen hindeutet. Insgesamt unterstreicht das Modell die enge Verknüpfung zwischen Zinsstruktur, Inflationserwartungen und Marktstimmung bei der Bestimmung langfristiger US-Renditen.

### 4.3 Fehlerkorrekturmodelle (ECM): Langfristige Gleichgewichtsbeziehungen

Nachdem die univariaten ARIMA-Modelle gezeigt haben, dass vergangene Renditebewegungen einen begrenzten, aber nachweisbaren Informationsgehalt für kurzfristige Prognosen besitzen, wird im folgenden Abschnitt untersucht, inwiefern makroökonomische Fundamentaldaten zusätzliche Erklärungs- und Prognosekraft bieten.

Der Engle-Granger-Kointegrationstest weist auf eine signifikante langfristige Beziehung zwischen der zehnjährigen (DGS10) und der zweijährigen (DGS2) US-Treasury-Rendite hin. Die geschätzte ADF-Statistik der Residuen beträgt  $-3,478$  bei einem p-Wert von  $0,044$  (siehe Anhang A6) und liegt damit unter dem kritischen Wert auf dem 5%-Signifikanzniveau. Damit wird die Nullhypothese fehlender Kointegration verworfen. Allerdings ist die Kointegration zwischen DGS10 und DGS2 nur grenzwertig signifikant ( $p = 0,044$ ) und könnte durch mögliche Strukturbrüche oder Regimewechsel im Zinsumfeld beeinträchtigt sein.

Die fünf ECM-Spezifikationen (siehe Kapitel 3.3) liefern folgende Ergebnisse:

**Tabelle 9:** Prognosegüte der Fehlerkorrekturmodelle (ECM-Klasse)

Modell	MAE	MSE	Cor	Hit Ratio	WHR
ECM_1	0,1917	0.0639	0.1271	0,542	0,544
ECM_2	0,1914	0.0617	0.2414	0,598	0,634
ECM_3	0,1889	0.0611	0.2429	0,576	0,593
ECM_4	0,2426	0.1247	0.0803	0,4803	0,545
ECM_5	0.2578	0.1481	0.0569	0.498	0.554
Manual Final	0,1934	0.0632	0.2676	0,563	0,567

*Note.* Eigene Darstellung. Die Tabelle zeigt die Out-of-Sample-Prognoseergebnisse der Fehlerkorrekturmodelle (ECM) für die 10-jährigen US-Staatsanleiherenditen. Bewertet wurden die Modelle anhand des Mean Absolute Error (MAE), Mean Squared Error (MSE), der Korrelation zwischen prognostizierten und tatsächlichen Renditeänderungen (Cor) sowie der Trefferquote (Hit Ratio) und der gewichteten Trefferquote (WHR). Geringere Fehlerwerte (MAE, MSE) und höhere Hit Ratios weisen auf eine bessere Prognosequalität hin.

Die Ergebnisse der geschätzten Error-Correction-Modelle (ECM) zeigen, dass die Kombination von kurzfristiger Dynamik und langfristiger Gleichgewichtsbeziehung zu einer Steigerung der Prognosequalität führt. Unter allen Modellen erzielt ECM\_3, das den Error-Correction-Term (ECT) mit den Aktienmarktrenditen (SP500) kombiniert, weiterhin den

niedrigsten mittleren absoluten Fehler (MAE = 0,1889) und verbessert damit die Prognosegüte um rund 28,4 % gegenüber der naiven Benchmark sowie um etwa 1,6 % gegenüber dem besten univariaten ARIMA-Modell.

Hinsichtlich der Richtungsprognose liefert ECM\_2, das zusätzlich DGS10, DGS2 und SP500 integriert, mit einer Weighted Hit Ratio von 63,4 % das beste Ergebnis. Dies unterstreicht die Bedeutung der Kombination von Zinsstruktur- und Marktvariablen bei der Modellierung der Renditedynamik.

Die Kombination aus Error-Correction-Term und fundamentalen Variablen in ECM\_2 liefert wertvolle Einblicke in die kurzfristige und langfristige Zinsdynamik. Der geschätzte Fehlerkorrekturterm zeigt mit einem Koeffizienten von  $\gamma = -0,021$  ( $p = 0,491$ ) zwar das erwartete negative Vorzeichen, ist jedoch statistisch insignifikant, was auf eine langsame oder unvollständige Rückkehr der Renditen zum langfristigen Gleichgewicht hinweist. Kurzfristig dominiert die Entwicklung des Aktienmarkts. Die Renditeveränderungen des S&P 500 (SP500\_r\_lag1,  $\beta = 0,015$ ;  $p = 0,002$ ) wirken signifikant positiv auf die Veränderung der zehnjährigen Treasury-Rendite, was auf eine starke Interaktion zwischen Risikoappetit und Kapitalmarktzinssätzen schließen lässt. Die detaillierten Schätzergebnisse für das ECM\_2, einschließlich der Regressionskoeffizienten, Signifikanzniveaus und Diagnostikstatistiken, sind im Anhang A7 dokumentiert. Weder die verzögerte Veränderung der zehnjährigen Rendite (d\_DGS10\_lag1,  $\beta = 0,174$ ;  $p = 0,153$ ) noch jene der zweijährigen Rendite (d\_DGS2\_lag1,  $\beta = -0,004$ ;  $p = 0,960$ ) zeigen signifikante Effekte.

Die weiterführenden Modelle (ECM\_4 – ECM\_5) mit automatischer Variablenselektion oder zusätzlichen Lags verschlechtern die Prognosequalität (MAE > 0,24), was auf Überparametrisierung und Informationsredundanz hinweist. Das Manual Final Model erreicht mit einem MAE von 0,1934 und einer WHR von 0,567 eine moderate, aber stabile Performance, bleibt jedoch hinter ECM\_2 und ECM\_3 zurück.

Diese Ergebnisse deuten darauf hin, dass kurzfristige Schwankungen der langfristigen Zinsen vor allem durch Finanzmarktentwicklungen getrieben werden, während geldpolitische und zinsstrukturelle Variablen in diesem Modell eine untergeordnete Rolle spielen.

#### **4.4 Hybride ARIMAX-Modelle: Integration von autoregressiven, fundamentalen und exogenen Dynamiken**

Die hybriden ARIMAX-Modelle verknüpfen die autoregressive und gleitende Durchschnittsstruktur der univariaten ARIMA-Modelle mit den fundamentalen Einflussgrößen der zuvor geschätzten Regressions- und Fehlerkorrekturmodelle. Durch diese Kombination wird sowohl die zeitliche Persistenz der Renditen als auch deren makroökonomische Determinierung berücksichtigt. Ziel dieser Modellklasse ist es, zu prüfen, ob die gleichzeitige Einbeziehung vergangener Renditeinformationen und fundamentaler Variablen die Prognosegüte weiter erhöht und somit ein umfassenderes

Abbild der Zinsdynamik liefert. Die folgende Tabelle 10 zeigt die Performance der ARIMAX Modelle:

**Tabelle 10:** Prognosegüte der ARIMAX-Modelle

Modell	<i>MAE</i>	MSE	Cor	Hit Ratio	WHR
ARIMAX	0.1888	0.0610	0.2529	0,598	0,608
ARIMAX Rol. 2J	0.2517	0.1036	-0.0152	0,493	0,503
ARIMAX Rol. 5J	0.2045	0.0728	0.0627	0,507	0,488
ARIMAX Rol. 10J	0.1962	0.0676	0.1264	0,542	0,529
ARIMAX Rol. 15J	0.1916	0.0659	0.1243	0,559	0,553
ARIMAX Rol. 20J	0.1926	0.0644	0.1600	0,572	0,570

*Note.* Eigene Darstellung. Die Tabelle zeigt die Out-of-Sample-Prognoseergebnisse der ARIMAX-Modelle, die sowohl autoregressive Komponenten als auch exogene makroökonomische Variablen integrieren. Bewertet wurden die Modelle anhand des Mean Absolute Error (MAE), Mean Squared Error (MSE), der Korrelation zwischen prognostizierten und tatsächlichen Renditeänderungen (Cor) sowie der Trefferquote (Hit Ratio) und der gewichteten Trefferquote (WHR). Geringere Fehlerwerte (MAE, MSE) und höhere Hit Ratios deuten auf eine bessere Prognosequalität hin.

Die Ergebnisse der hybriden ARIMAX-Modelle zeigen, dass die Kombination von Zeitreihendynamik und fundamentalen Einflussgrößen zu einer weiteren Verbesserung der Prognosequalität führt. Das ARIMAX-Basismodell, welches autoregressive Komponenten mit fundamentalen Variablen kombiniert, erzielt mit einem MAE von 0,1888 die beste Gesamtperformance dieser Modellklasse und liegt damit nahezu auf dem Niveau des leistungsstärksten ECM-Modells (ECM\_3: MAE = 0,1889). Besonders hervorzuheben ist das ARIMAX Rolling 20J, das mit einer Weighted Hit Ratio von 57,0 % die höchste Richtungsgenauigkeit innerhalb dieser Modellklasse erreicht. Das Modell zeigt damit, dass längere rollierende Schätzfenster eine stabile Balance zwischen Anpassungsfähigkeit und Prognosekonsistenz bieten.

Insgesamt reduzieren die besten ARIMAX-Spezifikationen den mittleren absoluten Fehler um rund 28,5 % und erhöhen die Weighted Hit Ratio um etwa 8 Prozentpunkte gegenüber der naiven Benchmark. Damit bestätigen die Modelle, dass die simultane Berücksichtigung vergangener Renditemuster und makroökonomischer Determinanten die Prognosefähigkeit der Zinsmodelle signifikant stärkt. Der Informationsgewinn durch die Integration fundamentaler Einflussgrößen zeigt zudem, dass Zinsbewegungen nicht ausschließlich

durch historische Renditemuster erklärbar sind, sondern in erheblichem Maße von makroökonomischen Rahmenbedingungen und Markterwartungen beeinflusst werden.

#### 4.5 Gesamtvergleich und Modellselektion

Die systematische Evaluation aller 27 geschätzten Modelle ermöglicht einen umfassenden Vergleich ihrer Prognosegüte. Tabelle 11 zeigt die zehn leistungsstärksten Modelle, sortiert nach dem mittleren absoluten Fehler (MAE). Neben den Fehlermaßen werden auch die Korrelationen zwischen prognostizierten und tatsächlichen Renditen berücksichtigt. Die vollständigen Ergebnisse aller Modelle sind im Anhang A8 dokumentiert.

**Tabelle 11:** Gesamtvergleich der Prognosegüte aller Modellklassen

Modell	MAE	MSE	Cor	Hit Ratio	WHR
ARIMAX	0.1888	0.0610	0.2529	0.5983	0.6083
ECM_3	0.1889	0.0611	0.2429	0.5764	0.5927
Modell_4	0.1908	0.0619	0.2123	0.5502	0.5619
ECM_2	0.1914	0.0617	0.2414	0.5983	0.6336
ARIMAX Rol. 15J	0.1916	0.0659	0.1243	0.5590	0.5532
ECM_1	0.1917	0.0639	0.1271	0.5415	0.5439
ARMA Rek. Opt.	0.1921	0.0631	0.1656	0.5590	0.5972
ARIMAX Rol. 20J	0.1926	0.0644	0.1600	0.5721	0.5699
Manual Final	0.1934	0.0647	0.0794	0.4891	0.5081
MA(1) Rekursiv	0.1934	0.0632	0.2676	0.5633	0.5671

*Note.* Eigene Darstellung. Die Tabelle fasst die Out-of-Sample-Prognoseergebnisse der besten Modelle aus allen untersuchten Klassen (ARIMA, ARIMAX, ECM und fundamentale Regressionsmodelle) zusammen. Bewertet wurden die Modelle anhand des Mean Absolute Error (MAE), Mean Squared Error (MSE), der Korrelation zwischen prognostizierten und tatsächlichen Renditeänderungen (Cor) sowie der Trefferquote (Hit Ratio) und der gewichteten Trefferquote (WHR). Geringere Fehlerwerte (MAE, MSE) und höhere Hit Ratios zeigen eine höhere Prognosequalität und Effizienz der jeweiligen Modellspezifikation.

Abbildung 3 veranschaulicht die Beziehung zwischen Prognosefehler (MAE) und Richtungsgenauigkeit (Weighted Hit Ratio) für die zehn besten Modelle. Es zeigt sich, dass Modelle mit geringeren Prognosefehlern tendenziell auch höhere Trefferquoten erzielen. Dies deutet darauf hin, dass präzisere Modelle nicht nur numerisch exakter, sondern auch

strukturell konsistenter im Verlauf ihrer Prognosen sind. Besonders hervorzuheben sind ECM\_2 und ARIMAX, die ein ausgewogenes Verhältnis zwischen Fehlerreduktion und Richtungsstabilität aufweisen und damit die effizientesten Modellansätze darstellen.

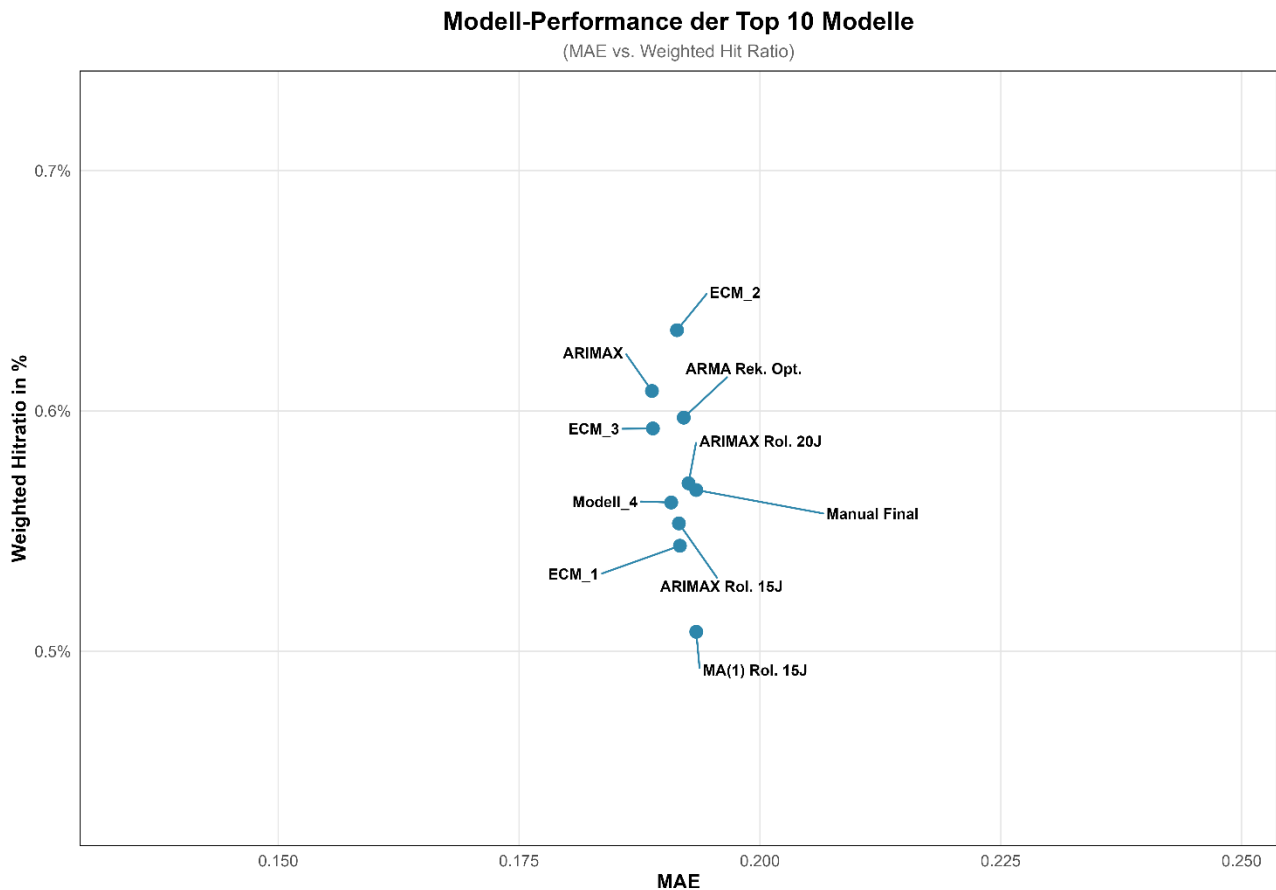


Abbildung 3: Darstellung der Modell-Performance der zehn besten Prognosemodelle (MAE vs. Weighted Hit Ratio); Quelle: eigene Darstellung. Die Abbildung zeigt die Beziehung zwischen Prognosefehler und Richtungsgenauigkeit für die zehn leistungsstärksten Modelle.

Das ARIMAX Modell bleibt das statistisch beste Modell mit dem niedrigsten MAE (0,1888) und der höchsten Korrelation ( $Cor = 0,2529$ ), was auf eine besonders enge lineare Übereinstimmung zwischen prognostizierten und tatsächlichen Renditebewegungen hinweist. Das ECM\_3 folgt mit nahezu identischer Prognosegüte (MAE = 0,1889) und einer stabilen Korrelation ( $Cor = 0,2429$ ) dicht dahinter. Hinsichtlich der Richtungsgenauigkeit erzielt das ECM\_2 die besten Ergebnisse mit einer Hit Ratio von 59,8 % und einer Weighted Hit Ratio von 63,4 %, womit es das stärkste Richtungsprognose-Modell darstellt. Diese Resultate belegen, dass sowohl die Integration langfristiger Gleichgewichtsbeziehungen als auch die Einbeziehung fundamentaler Variablen die Prognosefähigkeit der Modelle signifikant verbessert.

Die hybriden ARIMAX-Modelle (z. B. ARIMAX Rolling 15J und 20J) zeigen ebenfalls robuste Ergebnisse, erreichen jedoch geringere Korrelationen (0,12 – 0,16) und bleiben damit leicht hinter den ECM-Varianten zurück. Dies deutet darauf hin, dass die in den ECM-Residuen enthaltene Information bereits weite Teile der autoregressiven Dynamik abbildet. Auffällig

ist zudem, dass die univariaten ARMA-Modelle trotz geringerer Komplexität stabile Korrelationen (0,166) und geringe Fehlerwerte ( $\approx 0,192$ ) erzielen, was ihre methodische Robustheit bestätigt.

#### 4.6 Handelsstrategien: Evaluation der Profitabilität

Die vorangegangene Analyse zeigt, dass insbesondere das ARIMAX Modell (MAE = 0,1888; Cor = 0,2529) sowie die Error-Correction-Modelle (ECM 3: MAE = 0,1889; ECM 2: WHR = 63,4 %) die statistisch präzisesten Prognosen liefern. Diese Modelle kombinieren kurzfristige Dynamik mit fundamentalen Einflussgrößen und bilden damit die komplexe Zinsstruktur besonders realitätsnah ab. Aus praktischer Perspektive stellt sich jedoch die entscheidende Frage, ob diese Modelle auch ökonomisch verwertbare Informationen generieren, also in der Lage sind, profitablen Handelsstrategien gegenüber passiven Anlagestrategien einen Mehrwert zu verschaffen.

Zur Beantwortung dieser Frage wird im Folgenden die wirtschaftliche Nutzbarkeit der Prognosen anhand simulierter Long-Short-Handelsstrategien auf den 10-jährigen US-Treasury-Future untersucht. Dabei wird die risikoadjustierte Performance bewertet. Die Analyse erfolgt in drei Schritten: (1) Vergleich der ARIMA-Modelle untereinander, (2) Vergleich der fundamentalen und ECM-Modelle sowie (3) Gesamtvergleich inklusive der hybriden ARIMAX-Modelle.

Die Simulation der Handelsstrategien basiert auf vereinfachenden Annahmen, um den reinen Informationsgehalt der Prognosen isoliert zu bewerten. Es werden keine Transaktionskosten berücksichtigt, der Treasury-Future-Markt wird als vollkommen liquide angenommen, und es erfolgt kein Einsatz von Leverage, sodass alle Positionen vollständig eigenkapitalfinanziert sind.

Im ersten Schritt (1) werden die univariaten ARIMA- und MA-Modelle hinsichtlich ihrer Handelsperformance verglichen. Tabelle 12 fasst die Ergebnisse zusammen und zeigt, dass insbesondere mittelfristig geschätzte Modelle positive risikoadjustierte Renditen erzielen. Die Ergebnisse belegen, dass alle ARIMA-basierten Strategien mit mittleren Schätzfenstern die naive Buy-and-Hold-Strategie deutlich übertreffen.

**Tabelle 12:** Performancekennzahlen der modellbasierten Handelsstrategien

Modell	<i>Cumulative Return</i>	Ann. Return	Ann. Volatility	Sharpe Ratio	Max DD
ARMA Rek. Opt.	0.5215	0.0222	0.0579	0.3839	0.1714
ARMA RoI. Fest 20J	0.0772	0.0039	0.0583	0.0670	0.1633
MA(1) RoI. 10J	0.0529	0.0027	0.0583	0.0464	0.2058

MA(1) Rol. 20J	-0.0708	-0.0038	0.0583	-0.0658	0.1875
MA(1) Rol. 15J	-0.0778	-0.0042	0.0583	-0.0726	0.1936
MA(1) Rekursiv	-0.1127	-0.0062	0.0583	-0.1071	0.2224
Naive	-0.1959	-0.0114	0.0583	-0.1951	0.2777
MA(1) Rol. 5J	-0.2012	-0.0117	0.0583	-0.2009	0.2995
MA(1) Rol. 2J	-0.2236	-0.0132	0.0582	-0.2262	0.2852
ARMA Rol. Opt. 20J	-0.3044	-0.0188	0.0581	-0.3242	0.3744

*Note.* Eigene Darstellung. Die Tabelle zeigt die Ergebnisse der modellbasierten Long/Short-Handelsstrategien auf Basis der Prognosen der univariaten Modelle. Dargestellt sind der kumulierte Ertrag (Cumulative Return), die annualisierte Rendite (Ann. Return), die annualisierte Volatilität (Ann. Volatility), die Sharpe Ratio sowie der maximale Drawdown (Max DD). Höhere Sharpe Ratios und geringere Drawdowns weisen auf ein günstigeres Risiko-Rendite-Profil hin, während die annualisierte Rendite den durchschnittlichen jährlichen Ertrag der jeweiligen Strategie beschreibt.

Das ARMA rekursiv optimiert erzielt mit einer Sharpe Ratio von 0,38 und einer annualisierten Rendite von 2,22 % die beste Performance unter den univariaten Modellen. Gleichzeitig weist es mit einem maximalen Drawdown von –17,1 % ein günstiges Risiko-Rendite-Profil auf. Das ARMA Rolling Fest (20J) erreicht mit einer Sharpe Ratio von 0,07 und einer Rendite von 0,39 % eine leicht positive, aber deutlich schwächere risikoadjustierte Performance. Kürzere Schätzfenster, insbesondere das MA(1) Rolling 2J, führen hingegen zu negativen Sharpe Ratios (–0,23) und deutlich höheren Drawdowns (–28,5 %), was auf die Instabilität bei zu geringen Stichprobengrößen hinweist.

Aus Risiko-Sicht verbessert sich der maximale Drawdown von –27,8 % (Naive) auf –17,1 % (ARMA Rek. Opt.), ein Rückgang um fast 40 %. Dies bestätigt die höhere Stabilität der rekursiv optimierten Strategien gegenüber der Benchmark. Insgesamt zeigen die Ergebnisse, dass univariate Modelle mit stabilen, ausreichend langen Schätzfenstern eine robuste Grundlage für regelbasierte Handelsstrategien bieten.

Im zweiten Schritt (2) werden die fundamentalen Regressionsmodelle sowie die Error-Correction-Modelle (ECM) hinsichtlich ihrer Handelsperformance untersucht. Tabelle 13 fasst die Ergebnisse zusammen und zeigt, dass insbesondere Modelle mit langfristigen Gleichgewichtsbeziehungen und moderater Variablenselektion überdurchschnittliche risikoadjustierte Renditen erzielen, während überparametrisierte oder automatisch selektierte Modelle eine merklich schwächere Performance aufweisen.

**Tabelle 13:** Performancekennzahlen der fundamentalen und ECM-basierten Handelsstrategien

Modell	<i>Cumulative Return</i>	Ann. Return	Ann. Volatility	Sharpe Ratio	Max DD
ECM_2	1.0773	0.0391	0.0572	0.6831	0.0835
Modell_2	0.6524	0.0267	0.0578	0.4617	0.1676
ECM_3	0.5918	0.0247	0.0578	0.4264	0.1006
Manual Final	0.5793	0.0242	0.0579	0.4189	0.1676
Modell_4	0.5650	0.0237	0.0579	0.4104	0.1192
Modell_1	0.5333	0.0226	0.0579	0.3911	0.2031
ECM_5	0.3354	0.0153	0.0581	0.2628	0.2102
Modell_3	0.3309	0.0151	0.0581	0.2597	0.2103
Stepwise	0.2615	0.0122	0.0582	0.2105	0.2786
ECM_4	0.2615	0.0122	0.0582	0.2105	0.2786
ECM_1	0.1814	0.0088	0.0582	0.1506	0.1584

*Note.* Eigene Darstellung. Die Tabelle zeigt die Out-of-Sample-Ergebnisse der modellbasierten Long/Short-Handelsstrategien auf Basis der fundamentalen Regressions- und Fehlerkorrekturmodelle (ECM). Dargestellt sind der kumulierte Ertrag (Cumulative Return), die annualisierte Rendite (Ann. Return), die annualisierte Volatilität (Ann. Volatility), die Sharpe Ratio sowie der maximale Drawdown (Max DD). Höhere Sharpe Ratios und geringere Drawdowns deuten auf eine bessere risikoadjustierte Performance hin, während die annualisierte Rendite die durchschnittliche jährliche Wertentwicklung der jeweiligen Strategie widerspiegelt.

Das ECM\_2 weist die beste risikoadjustierte Performance unter allen getesteten fundamentalen und ECM-basierten Strategien auf. Mit einer Sharpe Ratio von 0,683, einer annualisierten Rendite von 3,91 % und einem maximalen Drawdown von 8,35 % erzielt es das stabilste und zugleich effizienteste Risiko-Rendite-Profil aller Modelle. Das Modell\_2 erreicht mit einer Sharpe Ratio von 0,462 und einer annualisierten Rendite von 2,67 % solide Ergebnisse, verzeichnet jedoch einen höheren Drawdown von 16,76 %. Das ECM\_3 erzielt mit einer Sharpe Ratio von 0,426 und einer Rendite von 2,47 % weiterhin gute Werte, bleibt jedoch klar hinter dem Spitzenmodell ECM\_2 zurück.

Das Manual Final (Sharpe = 0,419) und das Modell\_4 (Sharpe = 0,410) zeigen ein stabiles, wenn auch weniger effizientes Risiko-Rendite-Profil. Deutlich schwächer schneiden hingegen ECM\_5 und das Stepwise-Modell ab, deren Sharpe Ratios bei rund 0,26–0,27 liegen. Diese Ergebnisse unterstreichen, dass eine gezielte Variablenauswahl und eine

zurückhaltende Modellspezifikation entscheidend sind, um Überanpassung zu vermeiden und die Generalisierbarkeit der Modelle zu sichern.

Im Vergleich zu den besten univariaten Zeitreihenmodellen zeigt sich der Mehrwert der ECM-Ansätze deutlich. Gegenüber dem ARMA rekursiv optimiert verbessert das ECM\_2 die Sharpe Ratio um rund 78 % (0,68 vs. 0,38), die annualisierte Rendite um 76 % (3,91 % vs. 2,22 %) und reduziert den maximalen Drawdown um 51 % (−8,35 % vs. −17,14 %).

Im dritten Schritt (3) werden die ARIMAX-Hybridmodelle in den Gesamtvergleich integriert. Tabelle 14 fasst die Ergebnisse der zehn besten Strategien zusammen und zeigt, dass Modelle, die fundamentale Einflussgrößen dynamisch in die Zeitreihenstruktur einbinden, tendenziell die höchste risikoadjustierte Performance erzielen.

**Tabelle 14:** Vergleich der Performancekennzahlen aller Modellklassen

Modell	Kategorie	<i>Cumulative Return</i>	Ann. Return	Ann. Volatility	Sharpe Ratio	Max DD
ECM_2	Fundamental/E CM	1.0773	0.039	0.0572	0.6831	0.084
ARIMAX	ARIMAX	0.8201	0.032	0.0575	0.5541	0.096
Modell_2	Fundamental/E CM	0.6524	0.027	0.0578	0.4617	0.168
ECM_3	Fundamental/E CM	0.5918	0.025	0.0578	0.4264	0.101
Manual Final	Fundamental/E CM	0.5793	0.024	0.0579	0.4189	0.168
Modell_4	Fundamental/E CM	0.5650	0.024	0.0579	0.4104	0.120
Modell_1	Fundamental/E CM	0.5333	0.023	0.0579	0.3911	0.203
ARMA Rek. Opt.	ARIMA	0.5215	0.022	0.0579	0.3839	0.171
ARIMAX Rol. 20J	ARIMAX	0.4409	0.019	0.0580	0.3331	0.124
ECM_5	Fundamental/E CM	0.3354	0.015	0.0581	0.2628	0.210

*Note.* Eigene Darstellung. Die Tabelle zeigt die Out-of-Sample-Ergebnisse der modellbasierten Long/Short-Handelsstrategien für alle untersuchten Modellklassen (ARIMA, ARIMAX, fundamentale Regressionsmodelle und Fehlerkorrekturmodelle – ECM). Dargestellt sind der kumulierte Ertrag (Cumulative Return), die annualisierte Rendite (Ann. Return), die annualisierte Volatilität (Ann. Volatility), die Sharpe Ratio sowie der maximale Drawdown (Max DD). Höhere Sharpe Ratios und geringere Drawdowns kennzeichnen ein besseres Risiko-Rendite-Verhältnis, während die annualisierte Rendite den durchschnittlichen jährlichen Ertrag der jeweiligen Strategie widerspiegelt.

Die Ergebnisse des Gesamtvergleichs in Tabelle 14 zeigen, dass die besten Strategien weiterhin von den multivariaten beziehungsweise hybriden Modellklassen dominiert werden. An der Spitze steht das ECM\_2 mit einer Sharpe Ratio von 0,683, einer annualisierten Rendite von 3,91 % und einem maximalen Drawdown von 8,35 %. Auf dem zweiten Platz folgt das ARIMAX-Basismodell, das mit einer Sharpe Ratio von 0,555 und einer annualisierten Rendite von 3,20 % ebenfalls eine hohe risikoadjustierte Performance erzielt. Den dritten Rang belegt das Modell 2, dessen Sharpe Ratio von 0,462 und annualisierte Rendite von 2,67 % eine solide, wenn auch geringere Effizienz aufweisen.

Damit bestätigt sich das zentrale Ergebnis der Analyse. Modelle mit fundamentalen Einflussgrößen und langfristigen Gleichgewichtsmechanismen besitzen eine klar überlegene Prognose- und Handelsergebnisqualität. Alle Top-7-Strategien gehören zur Gruppe der multivariaten bzw. hybriden Modelle (ECM, Fundamental oder ARIMAX) und stützen somit die Hypothese, dass die Einbindung makroökonomischer Informationen und Kointegrationsbeziehungen zu einer ökonomisch verwertbaren Vorhersagekraft führt.

Besonders die ECM-Modelle belegen die Relevanz von Mean-Reversion-Mechanismen. Sie reagieren effizient auf Abweichungen vom langfristigen Gleichgewicht und liefern dadurch stabilere Risiko-Rendite-Profile. Das ARIMAX-Basismodell auf Rang 2 zeigt zudem, dass die Kombination aus autoregressiver Dynamik und fundamentalen Variablen zu einer robusten Verbesserung der Prognosegüte führt.

Im direkten Vergleich zur passiven Buy-and-Hold-Strategie zeigt sich, dass das ECM\_2 eine klar überlegene risikoadjustierte Performance erzielt. Während die Benchmark lediglich eine Sharpe Ratio von 0,021 und eine annualisierte Rendite von 0,12 % bei einem maximalen Drawdown von -24,48 % aufweist, erreicht das ECM\_2 eine Sharpe Ratio von 0,683, eine annualisierte Rendite von 3,91 % und einen maximalen Drawdown von -8,35 %. Damit übertrifft das Modell die naive Strategie um rund +3.150 % in der Sharpe Ratio, +3.150 % beim annualisierten Ertrag und reduziert gleichzeitig das maximale Verlustrisiko um -66 %. Diese Ergebnisse unterstreichen den erheblichen Mehrwert aktiver, modellbasierter Prognoseansätze gegenüber einem rein passiven Halten der Anleiheposition. Das ECM\_2 erzielt somit nicht nur höhere absolute Renditen, sondern verbessert auch das Risiko-Rendite-Profil signifikant, was auf eine effiziente Nutzung von Informations- und Gleichgewichtsdynamiken schließen lässt.

### **Risiko- und Portfolioanalyse**

Im Anschluss an den Gesamtvergleich der Strategien wird die Risiko-Rendite-Struktur der zehn besten Modelle analysiert. Ziel dieser Analyse ist es, die Stabilität und Effizienz der jeweiligen Strategien unter Berücksichtigung der Volatilität und der erzielten Renditen zu bewerten. Abbildung 4 stellt hierzu das Risiko-Rendite-Profil der Top-10-Modelle grafisch

dar und veranschaulicht, welche Modellklassen die beste Balance zwischen Risiko und Ertrag aufweisen.

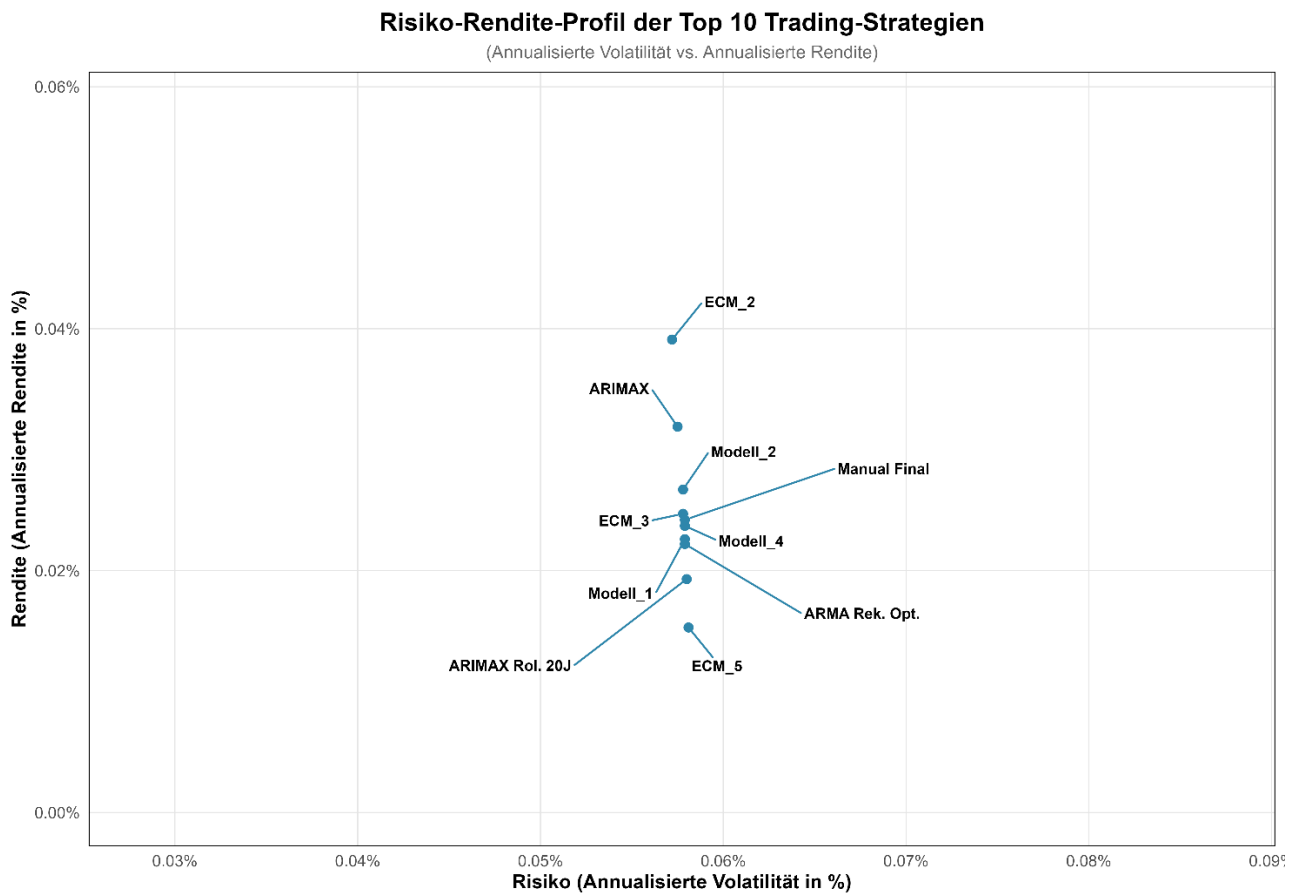


Abbildung 4: Risiko-Rendite-Profil der zehn besten Trading-Strategien (Annualisierte Volatilität vs. Annualisierte Rendite); Quelle: Eigene Darstellung auf Basis der Simulationsergebnisse (1977–2025); Datenbasis FRED (DGS10), eigene Berechnungen.

Die in Abbildung 4 dargestellte Risiko-Rendite-Analyse verdeutlicht, dass Modelle mit fundamentaler und hybrider Struktur auch im Risiko-Ertrags-Vergleich dominieren.

An der Spitze steht das ECM\_2, das bei einer annualisierten Rendite von 3,91 % und einer Volatilität von 5,72 % das effizienteste Risiko-Ertrags-Profil aufweist. Es kombiniert damit die höchste Stabilität mit dem größten Ertragspotenzial.

Das ARIMAX-Basismodell positioniert sich mit einer Rendite von 3,20 % und einer Volatilität von 5,75 % unmittelbar dahinter, was seine Rolle als robustes, dynamisches Hybridmodell bestätigt. Die Modelle 2 und 3 folgen mit leicht geringeren Renditen (2,67 % bzw. 2,47 %), zeigen jedoch eine vergleichbare Risikostruktur. Dies spricht für eine robuste Übertragbarkeit der fundamentalen Ansätze, auch wenn der absolute Ertrag etwas niedriger ausfällt.

Insgesamt verdeutlicht das Diagramm, dass Modelle mit langfristigen Gleichgewichtsbeziehungen (ECM) und fundamentaler Informationsintegration (ARIMAX)

nicht nur prognostisch überlegen sind, sondern auch aus Risiko-Ertrags-Perspektive die effizientesten Strategien darstellen.

### **Portfoliostrukturierung: Vergleich alternativer Allokationsansätze**

Auf Grundlage der zuvor ermittelten Modellperformances werden im Folgenden drei Portfolio-Strategien verglichen, die unterschiedliche Ansätze zur Diversifikation und Risikooptimierung verfolgen. Ziel ist es, die Effizienz eines rein modellbasierten Ansatzes (Single-Strategy) gegenüber diversifizierten Multi-Strategy-Portfolios zu bewerten.

Die einfachste Variante besteht in einer vollständigen Allokation in das beste Einzelmodell, ein Single-Strategy-Ansatz mit dem ECM\_2. Dieses erzielt mit einer Sharpe Ratio von 0,683, einer annualisierten Rendite von 3,91 % und einem maximalen Drawdown von –8,35 % die höchste risikoadjustierte Einzelleistung. Der Vorteil dieses Ansatzes liegt in der klaren Fokussierung auf das überlegene Modell; der Nachteil besteht in der vollständigen Konzentration auf eine einzige Quelle prognostischer Unsicherheit.

Reduktion des modellspezifischen Risikos wird im zweiten Ansatz, Diversifizierter Multi-Strategy-Ansatz (Equal Weight Portfolio), eine gleichgewichtete Kombination der zehn besten Strategien gebildet. Dieses Portfolio erreicht eine kumulative Rendite von 0,6290, eine annualisierte Rendite von 2,59 %, eine Volatilität von 3,85 %, eine Sharpe Ratio von 0,673 und einen maximalen Drawdown von –7,26 %. Im Vergleich zur Buy-and-Hold-Benchmark (Sharpe Ratio = 0,021; Drawdown = –24,48 %) zeigt sich eine deutliche Effizienzsteigerung, die den Diversifikationseffekt der Modellkombination bestätigt.

Der dritte Ansatz, Minimum-Variance-Portfolio (optimierte Gewichtung), optimiert die Gewichtung der Strategien nach dem Minimum-Variance-Kriterium. Dieses Portfolio erzielt die beste Gesamteffizienz mit einer Sharpe Ratio von 0,752, einer annualisierten Rendite von 2,37 %, einer Volatilität von 3,16 % und einem maximalen Drawdown von –5,63 %. Damit weist das Minimum-Variance-Portfolio die stabilste Performance im gesamten Untersuchungszeitraum auf und minimiert gleichzeitig die Schwankungsintensität gegenüber allen Vergleichsstrategien.

### **Optimierte Gewichte (Minimum Variance-Portfolio)**

**Tabelle 1515:** Optimierte Gewichte des Minimum-Variance-Portfolios

Modelle	<i>Gewichtung</i>
ARMA Rek. Opt.	0.3786
Modell_1	0.1535
ECM_3	0.1477
Manual Final	0.1181
ARIMAX Rol. 20J	0.1038

Modell_4	0.0416
ECM_5	0.0308
Modell_2	0.0257
ECM_2	0.0000
ARIMAX	0.0000

*Note.* Eigene Darstellung. Die Tabelle zeigt die im Rahmen der Minimum-Variance-Optimierung ermittelten Portfolioanteile der einzelnen Modelle. Ziel der Optimierung war die Minimierung der Portfoliovarianz bei gegebener erwarteter Rendite. Höhere Gewichtungen deuten auf eine überdurchschnittliche risikobereinigte Performance und eine stärkere Stabilität der Modellprognosen hin. Modelle mit einer Gewichtung von 0,0000 wurden aufgrund geringer Korrelation oder schwacher Sharpe Ratios nicht in das optimale Portfolio aufgenommen.

Die optimierten Gewichte zeigen, dass das Portfolio die Varianzreduktion vor allem durch eine Übergewichtung der Modelle ARMA Rek. Opt., Modell\_1 und ECM\_3 erzielt. Die Top-Modelle ECM\_2 und ARIMAX werden hingegen aufgrund ihrer höheren Korrelationen mit den übrigen Strategien im Optimierungsprozess nur gering gewichtet. Dies verdeutlicht, dass eine varianzbasierte Diversifikation nicht zwingend auf die renditestärksten Modelle setzt, sondern auch Strategien mit komplementären Risikoeigenschaften berücksichtigt, um das Gesamtrisiko effektiv zu minimieren.

Zur Beurteilung der in Kapitel 4 entwickelten Handels- und Portfolioansätze wurden zentrale Performancemaße wie die kumulative und annualisierte Rendite, die Sharpe Ratio, die annualisierte Volatilität sowie der maximale Drawdown herangezogen. Tabelle 16 fasst die Ergebnisse zusammen und ermöglicht einen direkten Vergleich zwischen den Top-5 modellbasierten Strategien, der Portfolio-Performance und der passiven Buy-and-Hold-Benchmark.

**Tabelle 1616:** Vergleich der Portfolio-Performance modellbasierter Strategien

Modell/ Portfolio	<i>Cumulative Return</i>	Ann. Return	Ann. Volatility	Sharpe Ratio	Max DD
Minimum Variance	0.5650	0.0237	0.0316	0.7523	0.0563
ECM_2	1.0773	0.0391	0.0572	0.6831	0.0835
Equal Weight	0.6290	0.0259	0.0385	0.6729	0.0726
ARIMAX	0.8201	0.032	0.0575	0.5541	0.096
Modell_2	0.6524	0.0267	0.0578	0.4617	0.1676

ECM_3	0.5918	0.0247	0.0578	0.4264	0.1006
Manual Final	0.5793	0.0242	0.0579	0.4189	0.1676
Buy & Hold	0.0233	0.0012	0.0583	0.0207	0.2448

*Note.* Eigene Darstellung. Die Tabelle vergleicht die Out-of-Sample-Performance verschiedener modellbasierter Portfolios (z. B. ECM-, Regressions- und Benchmarkstrategien) mit alternativen Allokationsansätzen. Dargestellt sind der kumulierte Ertrag (Cumulative Return), die annualisierte Rendite (Ann. Return), die annualisierte Volatilität (Ann. Volatility), die Sharpe Ratio sowie der maximale Drawdown (Max DD). Höhere Sharpe Ratios und geringere Drawdowns kennzeichnen ein günstigeres Risiko-Rendite-Profil, während die annualisierte Rendite den durchschnittlichen jährlichen Ertrag der jeweiligen Strategie widerspiegelt.

Deskriptiv zeigt sich, dass insbesondere das ECM\_2 mit einer annualisierten Rendite von 3,91 % und einer Sharpe Ratio von 0,683 die höchste risikoadjustierte Einzelleistung erzielt. Darüber hinaus weisen die diversifizierten Portfolios eine nochmals verbesserte Effizienz auf. Das Minimum-Variance-Portfolio erreicht mit einer Sharpe Ratio von 0,752 und einem maximalen Drawdown von 5,63 % die stabilste Gesamtpformance, gefolgt vom gleichgewichteten Portfolio mit einer Sharpe Ratio von 0,673.

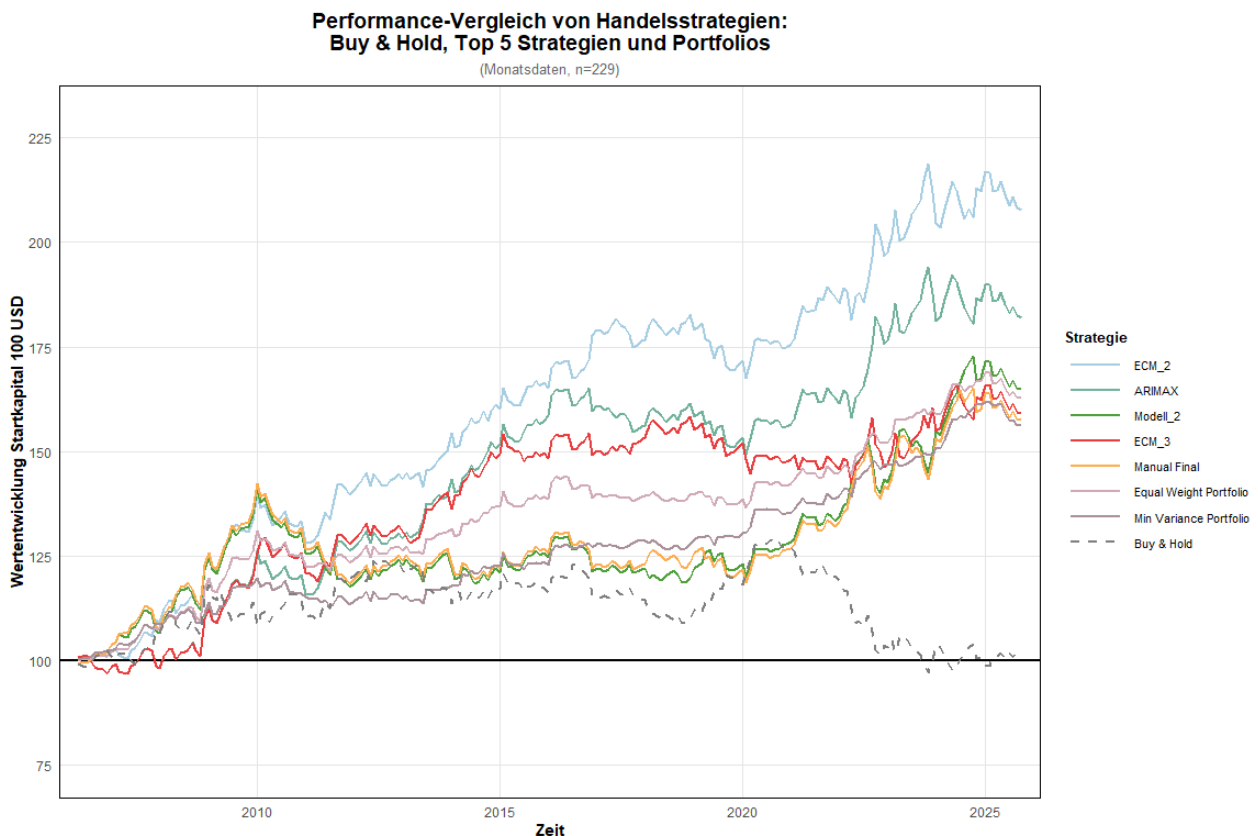


Abbildung 5: Performance-Vergleich der modellbasierten Handelsstrategien im Zeitraum 2006–2025 auf Monatsbasis. Dargestellt sind die Wertentwicklungen eines Startkapitals von 100 USD für die besten Einzelmodelle (ECM\_2, ARIMAX, Modell\_2, ECM\_3, Manual Final) sowie die beiden diversifizierten Portfolios (Equal Weight Portfolio und Minimum Variance Portfolio) im Vergleich zur passiven Buy-and-Hold-Strategie

auf 10-jährige US-Treasuries; Quelle: Eigene Darstellung auf Basis der Simulationsergebnisse (1977–2025); Datenbasis FRED, eigene Berechnungen.

Die grafische Gegenüberstellung der Handelsstrategien in Abbildung 5 verdeutlicht die deutlichen Leistungsunterschiede zwischen aktiven, modellgestützten Ansätzen und der passiven Benchmark. Während das Buy-and-Hold-Exposure auf 10-jährige US-Treasuries über den gesamten Beobachtungszeitraum lediglich einen Endwert von 102,33 USD (Rendite 2,33 %) erzielt und damit kaum Wertzuwachs generiert, zeigen alle modellbasierten Strategien eine deutlich höhere Kapitalentwicklung.

Das ECM\_2 weist mit einem Endwert von 207,73 USD und einer Gesamtrendite von 107,73 % die mit Abstand stärkste Performance auf. Die Strategie profitiert sichtbar von der Mean-Reversion-Dynamik der Zinsstruktur und reagiert adaptiv auf temporäre Fehlbewertungen am Markt. Auch das ARIMAX-Modell (182,01 USD / 82,01 %) und das Modell\_2 (165,24 USD / 65,24 %) liefern überdurchschnittliche Ergebnisse, was die Relevanz makroökonomischer Informationsintegration unterstreicht.

Die beiden kombinierten Portfolios erzielen ein stabiles, wenngleich etwas niedrigeres Niveau der Gesamtrendite: Das Equal Weight Portfolio erreicht einen Endwert von 162,90 USD, das Minimum Variance Portfolio von 156,50 USD. Trotz der geringeren absoluten Performance überzeugen beide durch ihre ausgeprägte Glättung der Wertentwicklung und die geringste Volatilität unter allen getesteten Strategien. Interessant ist, dass die Portfolios insbesondere während volatiler Marktphasen, etwa zwischen 2018 und 2020, deutliche Stabilitätsvorteile zeigen, was auf den Diversifikationseffekt der schwach korrelierten Modellprognosen zurückzuführen ist.

Die passive Buy-and-Hold-Strategie schneidet mit einer annualisierten Rendite von 0,12 %, einer Sharpe Ratio von 0,021 und einem maximalen Drawdown von 24,48 % am schwächsten ab. Damit bestätigt sich, dass ökonometrische Prognosemodelle, insbesondere die auf fundamentalen und kointegrativen Beziehungen basierenden ECM- und ARIMAX-Ansätze, nicht nur statistisch signifikante, sondern auch ökonomisch verwertbare Ergebnisse liefern und passiven Strategien in allen untersuchten Dimensionen klar überlegen sind.

## 5 Diskussion und Fazit

In diesem Kapitel werden die im vorangegangenen Abschnitt dargestellten empirischen Ergebnisse zusammenfassend bewertet, theoretisch eingeordnet und kritisch reflektiert. Ziel ist es, die Prognosegüte der untersuchten Modellklassen im Kontext der Forschungsfragen zu analysieren und daraus Schlussfolgerungen für Theorie und Praxis abzuleiten. Dabei werden sowohl die ökonometrische Leistungsfähigkeit der Modelle als auch deren praktische Relevanz im Rahmen der abgeleiteten Handelsstrategien und Portfolioallokationen betrachtet.

### 5.1 Zusammenfassung der zentralen Ergebnisse

Die Ergebnisse zeigen, dass multivariate und kointegrationsbasierte Modellansätze den univariaten Zeitreihenmodellen in nahezu allen Bewertungskriterien überlegen sind. Insbesondere das ECM\_2 weist mit einer Sharpe Ratio von 0,683, einer annualisierten Rendite von 3,91 % und einem maximalen Drawdown von 8,35 % die beste risikoadjustierte Performance aller getesteten Modelle auf. Damit weist das Modell das insgesamt effizienteste Rendite-Risiko-Profil auf und verdeutlicht den Mehrwert eines expliziten Fehlerkorrekturmechanismus zur Modellierung langfristiger Zinsdynamiken.

Im Vergleich dazu liefert das passive Buy-and-Hold-Exposure auf 10-jährige US-Treasuries lediglich eine durchschnittliche Rendite von 0,12 % p. a. bei einem Drawdown von 24,48 %, was die Überlegenheit aktiver, modellgestützter Strategien weiterhin deutlich macht.

Auch innerhalb der ökonometrischen Modelle zeigt sich eine klare Rangordnung. Während die univariaten ARIMA-Modelle stabile Grundlagen für kurzfristige Prognosen bieten, können sie die komplexen Zusammenhänge zwischen Zinsniveau, Inflation, Geldpolitik und Markterwartungen nur eingeschränkt abbilden.

Die Integration exogener Informationen über ARIMAX-Modelle führt zu spürbar besseren Ergebnissen hinsichtlich der Sharpe Ratio (0,55) und des Drawdowns (9,6 %), was den Informationsgehalt fundamentaler Variablen für Zinsprognosen unterstreicht. Darüber hinaus erreichen die hybriden ECM-Modelle, die kurz- und langfristige Dynamiken über den Error-Correction-Term verknüpfen, eine besonders hohe Stabilität und eine konsistent positive Performance über verschiedene Zeithorizonte und Spezifikationen hinweg. Die Resultate zeigen damit, dass Zinsmärkte keineswegs vollständig zufällig agieren, sondern auf nachhaltige Fundamentalmuster reagieren, die ökonometrisch quantifizierbar sind.

Von besonderem Interesse ist die Analyse der auf den Modellen basierten Handelsstrategien. Hier zeigt sich, dass die auf dem ECM\_2 basierende Long/Short-Strategie nicht nur eine überlegene Rendite, sondern auch eine geringere Risikobelastung aufweist. Die Ergebnisse deuten darauf hin, dass die Einbeziehung des Fehlerkorrektur-Terms als Signal für Rückkehrbewegungen in Renditen eine effektive Risikosteuerung ermöglicht.

Durch den Vergleich mit anderen ECM-Varianten zeigt sich zudem, dass Modelle mit zu vielen Freiheitsgraden tendenziell eine höhere Volatilität aufweisen und daher weniger robust sind. Ein weiterer zentraler Befund besteht darin, dass die Minimum-Variance-Kombination der leistungsstärksten Strategien, bei Korrelationen zwischen 0,35 und 0,55, die Sharpe Ratio auf 0,752 erhöht und die Volatilität auf 3,16 % senkt. Damit lässt sich zeigen, dass eine Diversifikation über mehrere schwach korrelierte Modelle zu einer effizienteren Rendite-Risikostruktur führt als die Nutzung eines einzelnen Modells.

In Summe lassen die Ergebnisse den Schluss zu, dass ökonomische und kointegrationsbasierte Modelle nicht nur theoretisch fundierter, sondern auch empirisch überlegen sind. Sie bieten eine präzisere Abbildung langfristiger Gleichgewichtsbeziehungen und eine effizientere Risikosteuerung als klassische zeitreihenbasierte Ansätze.

## 5.2 Beantwortung der Forschungsfragen

Die erste Forschungsfrage bezieht sich auf die Prognosefähigkeit unterschiedlicher Modellklassen und untersucht, welche Ansätze die höchste out-of-sample Prognosegüte für die 10-jährige US-Treasury-Rendite erzielen. Die Analyse der Ergebnisse zeigt, dass die Integration makroökonomischer Einflussgrößen zu einer signifikanten Verbesserung der Prognosequalität führt. Während univariate Zeitreihenmodelle wie die ARIMA-Klasse kurzfristige Renditedynamiken zuverlässig abbilden, bleiben sie bei der Erklärung struktureller Veränderungen im Zinsniveau begrenzt. Multivariate und hybride Modelle, insbesondere die ECM- und ARIMAX-Varianten, erreichen geringere Prognosefehler gemessen am Mean Absolute Error (MAE) und weisen zugleich höhere Werte für die Hit Ratio (HR) und die Weighted Hit Ratio (WHR) auf. Das ECM\_2 erweist sich als das leistungsstärkste Einzelmodell und erzielt mit einem MAE von 0,1914, einer Hit Ratio von 0,598 und einer Weighted Hit Ratio von 0,634 die insgesamt beste Kombination aus Prognosepräzision und Stabilität. Der im Modell enthaltene Fehlerkorrektur-Term (ECT) ermöglicht eine Anpassung kurzfristiger Renditebewegungen an das langfristige Gleichgewicht und reduziert systematische Prognoseabweichungen.

Die zweite Forschungsfrage widmet sich der praktischen Nutzbarkeit der Modellprognosen und untersucht, inwieweit sich diese in einfache Handelsstrategien mit überlegener risikoadjustierter Rendite übersetzen lassen. Die empirischen Ergebnisse zeigen, dass insbesondere das auf dem ECM\_2 basierende Long/Short-Handelssystem eine höhere Performance als der passive Buy-and-Hold-Ansatz erzielt. Mit einer Sharpe Ratio von 0,683, einer annualisierten Rendite von 3,91 % und einem maximalen Drawdown von 8,35 % weist die Strategie das beste Risiko-Ertrags-Profil aller getesteten Modelle auf. Zum Vergleich erzielt die passive Benchmark auf 10-jährige Treasuries lediglich eine Rendite von 0,12 % pro Jahr bei einem maximalen Drawdown von 24,48 %. Diese Differenz verdeutlicht den ökonomischen Mehrwert der modellbasierten Signale. Die Ergebnisse implizieren, dass Marktteilnehmer, die strukturierte Prognosen auf Basis fundamentaler Informationen

nutzen, in der Lage sind, zeitweise Ineffizienzen auszunutzen und so eine höhere risikobereinigte Rendite zu erzielen, ohne ein übermäßiges Risiko einzugehen.

Die dritte Forschungsfrage untersucht, ob sich durch die Kombination mehrerer Modellstrategien in einem diversifizierten Portfolio eine zusätzliche Verbesserung der Stabilität und Performance erzielen lässt. Die Ergebnisse der Portfolioanalyse zeigen, dass die Diversifikation über schwach korrelierte Modellansätze substantielle Effizienzgewinne ermöglicht. Das Minimum-Variance-Portfolio welches durch ARIMA (rekursiv optimiert, 0,38), Modell 1 (0,15), ECM 3 (0,15) und Manual Final (0,12) getragen wird, steigert die Sharpe Ratio auf 0,752 bei einer Volatilität von 3,16 %. Diese Verbesserung beruht auf der geringen Korrelation der Einzelsignale zwischen 0,35 und 0,55, die eine wirksame Glättung der Ertragsverteilung bewirkt. Damit liefert die Kombination der Prognosemodelle nicht nur eine höhere risikoadjustierte Rendite, sondern auch eine ausgeprägte Robustheit gegenüber einzelnen Modellschwächen. Die Resultate bestätigen den in der Literatur vielfach dokumentierten Nutzen von Forecast Combinations und zeigen, dass sich Diversifikation nicht nur auf Anlageklassen, sondern auch auf Informationsquellen und Modellarchitekturen anwenden lässt (Armstrong, 2001, ch. 13).

### **5.3 Theoretische Einordnung und Implikationen für die Effizienzmarkthypothese**

Die Ergebnisse dieser Arbeit stehen in engem Zusammenhang mit der Erwartungstheorie der Zinsstruktur, wonach langfristige Renditen als Funktion erwarteter zukünftiger Kurzfristzinsen und einer Termprämie modelliert werden. Der im ECM implementierte Fehlerkorrektur-Mechanismus sorgt dafür, dass Abweichungen vom langfristigen Gleichgewicht korrigiert werden und kurzfristige Renditebewegungen wieder in Richtung des theoretischen Gleichgewichts verlaufen. Damit stützen die empirischen Resultate die Annahme, dass Zinsentwicklungen nicht zufällig, sondern durch fundamentale makroökonomische Variablen erklärbar sind.

Das insbesondere ECM- und ARIMAX-Modelle die höchste Prognosegüte erzielen, weist auf eine unvollständige und zeitlich gestaffelte Informationsverarbeitung am US-Zinsmarkt hin. Diese Beobachtung steht im Einklang mit der Annahme einer adaptiven Markteffizienz, bei der Informationen zwar verarbeitet, jedoch nicht sofort vollständig in die Preise integriert werden. Folglich ergibt sich keine vollständige Zurückweisung der Markteffizienz, jedoch deutliche Evidenz gegen ihre starke Form. Wären sämtliche Informationen bereits vollständig eingepreist, könnten systematisch überlegene Prognosen wie jene des ECM\_2 nicht erzielt werden.

Die Ergebnisse verdeutlichen somit, dass die Integration makroökonomischer Variablen wie Inflation, Term Spread und Arbeitsmarktindikatoren einen realen Informationsgewinn liefert. Märkte reagieren adaptiv und mit zeitlicher Verzögerung auf fundamentale Veränderungen, wodurch vorübergehende Ineffizienzen entstehen. In Summe stehen die Ergebnisse im

Widerspruch zur semi-starken Effizienzmarkthypothese, bei der öffentlich verfügbare Informationen zwar relevant, aber nicht augenblicklich vollständig verarbeitet werden.

#### **5.4 Kritische Würdigung und Limitationen**

Trotz der robusten Ergebnisse und der theoretischen Fundierung unterliegt diese Untersuchung mehreren Limitationen, die bei der Interpretation der Resultate berücksichtigt werden müssen.

Ein wesentliches Risiko liegt in der Datenbasis und Modellstruktur. Die Ergebnisse hängen stark von der Auswahl der verwendeten Variablen, der Datenfrequenz und der Modellspezifikation ab. Die Analyse basiert auf monatlichen Beobachtungen, wodurch kurzfristige Marktreaktionen und hochfrequente Informationsanpassungen nur eingeschränkt abgebildet werden können. Zwar bieten Monatsdaten eine gute Balance zwischen Signal und Rauschen, dennoch könnten Modelle auf Wochen- oder Tagesbasis zusätzliche kurzfristige Dynamiken offenlegen. Die verwendeten makroökonomischen Indikatoren wie Inflation, Term Spread und Arbeitsmarktvariablen sind aggregierte Größen, deren Veröffentlichung zeitverzögert erfolgt. Damit reflektieren die Modelle reale Informationsprozesse nur approximativ. Zudem könnte die Wahl alternativer Variablensets oder Lag-Strukturen zu leicht verschobenen Ergebnissen führen, was die Robustheit einzelner Spezifikationen relativiert.

Ein weiterer limitierender Faktor betrifft die Stabilität der Modellparameter über verschiedene Zinsregime hinweg. Der Untersuchungszeitraum von 2006 bis 2025 umfasst mehrere außergewöhnliche Marktphasen, darunter die globale Finanzkrise 2008, die COVID-19-Pandemie 2020 sowie die Zinswende der Federal Reserve ab 2022. Solche Regimewechsel können strukturelle Brüche und Parameterdrift verursachen. Zwar reduzieren rollierende und expandierende Schätzverfahren die Auswirkungen solcher Instabilitäten, sie können sie jedoch nicht vollständig eliminieren. Zukünftige Arbeiten könnten hier Time-Varying-Parameter- oder Regime-Switching-Modelle einsetzen, um adaptive Anpassungsmechanismen genauer zu erfassen (Cordeiro et al., o.J.; Duffee, 2012).

Darüber hinaus bestehen methodische Einschränkungen der Handelsauswertung. Die simulierten Strategien berücksichtigen keine Transaktionskosten, Slippage oder Kommissionen. Unter realen Marktbedingungen würden diese Faktoren die Nettorendite um schätzungsweise 0,3 bis 0,5 Prozentpunkte pro Jahr reduzieren, was die Sharpe Ratio geringfügig absenken, die relative Überlegenheit gegenüber Buy-and-Hold jedoch nicht aufheben würde (DeMiguel, Garlappi & Uppal, 2009). Ebenso setzt die Simulation eine perfekte Marktliquidität voraus, was in der Praxis nur näherungsweise gilt. Zwar ist der US-Treasury-Future-Markt einer der liquidesten der Welt (Chaboud et al., 2025), dennoch können große Orders den Markt kurzfristig beeinflussen und die Umsetzungskosten erhöhen. Auch wurde mit fixen Positionsgrößen gearbeitet, wodurch dynamische

Anpassungen an Volatilität oder Signalstärke unberücksichtigt bleiben. Ein Volatility-Targeting- oder Kelly-basiertes Positionsmanagement könnte die risikoadjustierte Performance künftig weiter verbessern (Cordeiro et al., o.J.).

Ein weiteres potenzielles Problem betrifft das Overfitting-Risiko (Rocha et al., 2024, ch. 4.3). Die Vielzahl an möglichen Variablen- und Lag-Kombinationen birgt die Gefahr, dass Modelle spezifische Muster der Trainingsperiode zu stark abbilden und ihre Generalisierbarkeit verlieren. Um dem entgegenzuwirken, wurden Out-of-Sample-Tests, rollierende Schätzfenster und Informationskriterien wie AIC und BIC zur Modellselektion eingesetzt. Die konsistente Überlegenheit der besten Modelle über verschiedene Teilperioden deutet darauf hin, dass die Resultate nicht auf zufälligen Dateneffekten beruhen, vollständig ausschließen lässt sich eine Überanpassung jedoch nicht.

Schließlich ist auch die Übertragbarkeit der Ergebnisse begrenzt. Die Untersuchung bezieht sich ausschließlich auf die 10-jährige US-Treasury-Rendite. Ob sich die Befunde auf andere Laufzeiten, Länder oder Assetklassen verallgemeinern lassen, bleibt offen. Unterschiede in Marktstruktur, Liquidität und geldpolitischem Regime könnten zu abweichenden Ergebnissen führen und sollten Gegenstand zukünftiger Forschung sein.

Insgesamt zeigen die genannten Einschränkungen, dass die hier präsentierten Resultate trotz hoher statistischer und ökonomischer Robustheit innerhalb eines klar umrissenen Analyserahmens zu interpretieren sind. Sie schmälern jedoch nicht die Aussagekraft der zentralen Erkenntnis, dass ökonomisch fundierte Modelle in der Lage sind, Informationsverzögerungen an Zinsmärkten zu identifizieren und in prognostische sowie strategische Vorteile zu übersetzen.

## 5.5 Ausblick für Forschung und Praxis

Die Ergebnisse und Limitationen dieser Arbeit eröffnen mehrere Anknüpfungspunkte für weiterführende Forschung und praxisorientierte Anwendungen. Auf wissenschaftlicher Ebene besteht zunächst ein erheblicher Mehrwert darin, die hier eingesetzten linearen Modelle um nicht-lineare und maschinelle Lernverfahren zu erweitern. Ansätze wie rekurrente neuronale Netze (LSTM), Gradient-Boosting-Verfahren oder Ensemble-Methoden wie Random Forests und Stacking könnten komplexe Interaktionen und zeitliche Abhängigkeiten besser erfassen als klassische ökonometrische Modelle (Fischer & Krauss, 2018; Lemke & Gabrys, 2010., Hall & Rasheed, 2025). Besonders interessant wäre die Untersuchung, ob solche Verfahren die Weighted Hit Ratio von 63,4 Prozent, die in dieser Arbeit für das ECM\_2 erreicht wurde, weiter verbessern können.

Darüber hinaus bietet sich eine Ausweitung der Analyse auf andere Laufzeiten und Märkte an. Die Berücksichtigung von kurz- und langfristigen US-Treasuries, Unternehmensanleihen oder internationalen Rentenmärkten könnte zeigen, ob die identifizierten Zusammenhänge, insbesondere Kointegration und Mean-Reversion, universell oder spezifisch für den US-Markt Zinsmarkt sind. Eine solche Erweiterung würde zugleich ermöglichen, die gesamte

Zinsstrukturkurve modellbasiert zu erfassen und dadurch umfassendere Aussagen über die Transmission geldpolitischer Impulse zu treffen.

Ein weiterer Ansatzpunkt liegt im Einbezug alternativer Datenquellen. Die Integration von Stimmungsindikatoren aus Finanznachrichten, Social-Media-Daten oder Textanalysen von FOMC-Statements könnte dazu beitragen, die kurzfristige Prognosequalität zu erhöhen, insbesondere in Phasen erhöhter Unsicherheit wie Zinswenden oder Krisenperioden. Ebenso könnten Hochfrequenzdaten, etwa aus Orderbüchern oder Tick-Daten, wertvolle Informationen über Marktliquidität und Mikrostruktur liefern. Der Einsatz solcher Datenarten würde es ermöglichen, sowohl makroökonomische als auch mikrodynamische Komponenten der Zinsbildung simultan zu erfassen.

Auch methodisch ergeben sich relevante Anschlussmöglichkeiten. Die Implementierung von Regime-Switching- oder Time-Varying-Parameter-Modellen könnte die Anpassungsfähigkeit der Modelle an Strukturbrüche weiter erhöhen. Insbesondere Markov-Switching-Modelle und TVP-VAR-Spezifikationen bieten die Möglichkeit, zwischen Phasen hoher und niedriger Volatilität zu unterscheiden und zeitvariante Beziehungen zwischen Fundamentalfaktoren und Renditen explizit abzubilden (Tan & Wu, 2025). Darüber hinaus erscheint es sinnvoll, die hier betrachtete Ein-Monats-Prognose um längere Horizonte von drei, sechs oder zwölf Monaten zu ergänzen, um zu prüfen, ob fundamentale Modelle mit zunehmendem Zeithorizont relativ an Prognosekraft gewinnen.

Neben diesen forschungsbezogenen Perspektiven ergeben sich auch konkrete Implikationen für die Praxis. Für Portfoliomanager und institutionelle Investoren zeigen die Ergebnisse, dass eine systematische quantitative Modellierung eine wertvolle Ergänzung traditioneller Anlageprozesse darstellt. Prognosen aus ECM-Modellen können zur taktischen Steuerung der Duration in Fixed-Income-Portfolios genutzt werden, indem erwartete Renditeanstiege eine Verkürzung und erwartete Rückgänge eine Verlängerung der Duration nahelegen. Gleichzeitig bieten Modelle wie beispielsweise das ECM\_2 oder ARIMAX, mit einer Weighted Hit Ratio von über 60 Prozent, eine solide Grundlage für ein präziseres Risikomanagement.

Für Zentralbanken und Aufsichtsbehörden liefern die identifizierten fundamentalen Treiber wichtige Anhaltspunkte für die Bewertung geldpolitischer Transmissionsmechanismen. Die beobachtete Kointegration zwischen kurz- und langfristigen Zinssätzen kann als Indikator für die Wirksamkeit geldpolitischer Maßnahmen interpretiert werden. Darüber hinaus könnte sie zur Überwachung der Finanzstabilität herangezogen werden, da eine Abkopplung der Renditestruktur auf mögliche Spannungen im Finanzsystem hinweisen kann.

Schließlich lassen sich auch Implikationen für den quantitativen Investmentbereich ableiten. Die gezeigte Outperformance der modellbasierten Strategien, mit einer Sharpe Ratio von 0,673 und einer nochmals gesteigerten 0,753 im Minimum-Variance-Portfolio, verdeutlicht das Potenzial einer diversifizierten Kombination mehrerer Modellarchitekturen. Hedge Funds und quantitative Vermögensverwalter könnten auf dieser Basis Multi-Strategy-

Ansätze entwickeln, bei denen unterschiedliche Prognosemodelle mit dynamischem Positionsmanagement kombiniert werden. Die Implementierung höherfrequenter Handelsansätze, beispielsweise auf Wochen- oder Tagesbasis, könnte dabei zusätzliche Alpha-Quellen erschließen.

Trotz der genannten Limitationen stellt die vorliegende Arbeit einen wichtigen Beitrag zur empirischen Erforschung von Zinsprognosen dar. Sie liefert belastbare Ergebnisse zur Prognosegüte und ökonomischen Verwertbarkeit verschiedener Modellklassen und bietet damit eine solide Grundlage für künftige Untersuchungen im Bereich der empirischen Zinsstrukturmodellierung. Die Ergebnisse zeigen, dass trotz der inhärenten Unsicherheiten an den Kapitalmärkten präzise und ökonomisch verwertbare Prognosen möglich sind und dass modellbasierte Strategien einen signifikanten Mehrwert gegenüber rein passiven Ansätzen bieten.

Im Spannungsfeld zwischen theoretischer Markteffizienz und praktischer Prognostizierbarkeit zeigt die Arbeit, dass Zinsmärkte zwar weitgehend effizient agieren, jedoch zeitweise systematische Abweichungen vom Gleichgewicht aufweisen. Diese Phasen eröffnen Spielräume für datengetriebene Modelle, die Informationsverzögerungen erkennen und in ökonomische Entscheidungsprozesse übersetzen können. Die Kombination aus fundamentaler Theorie und empirischer Modellierung ermöglicht somit einen ausgewogenen Blick auf die Funktionsweise moderner Finanzmärkte, in denen Rationalität, Erwartungen und adaptive Anpassungsprozesse nebeneinander bestehen.

Mit Blick auf die zentrale Bedeutung der 10-jährigen US-Staatsanleihe als globalen Referenzwert unterstreicht die Untersuchung ihre Schlüsselrolle im internationalen Finanzsystem. Ihre Rendite reflektiert nicht nur die geldpolitischen Erwartungen gegenüber der Federal Reserve, sondern dient zugleich als Gradmesser für Kapitalströme, Risikoaversion und makroökonomische Stabilität. Die in dieser Arbeit identifizierten Zusammenhänge zwischen fundamentalen Variablen und langfristigen Zinsen tragen dazu bei, diese Dynamiken ökonometrisch zu erfassen und besser zu verstehen. Vor dem Hintergrund der bevorstehenden geldpolitischen Entscheidungen der Federal Reserve und der Diskussionen über mögliche Zinssenkungen noch im Jahr 2025 gewinnt die zukünftige Entwicklung der langfristigen US-Zinsen erneut an Relevanz. Die vorgestellten Modelle bieten eine fundierte empirische Grundlage, um Veränderungen in der geldpolitischen Ausrichtung und deren Übertragung auf die Kapitalmärkte systematisch zu analysieren und ökonomisch einzuordnen.

## Literaturverzeichnis

- Ang, A., & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4), 745–787. Available at: [https://doi.org/10.1016/S0304-3932\(03\)00032-1](https://doi.org/10.1016/S0304-3932(03)00032-1)
- Armstrong, J. S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners (1st ed. 2001)*. Imprint: Springer. <https://doi.org/10.1007/978-0-306-47630-3>
- Bailey, D. H., Borwein, J. M., de Prado, M. L., & Zhu, Q. J. (2015). The probability of backtest overfitting. *Journal of Computational Finance*, 20(4), 39–69. <https://dx.doi.org/10.2139/ssrn.2326253>
- Bandholz, H., Clostermann, J., & Seitz, F. (2007). *Explaining the US bond yield conundrum*. *Applied Financial Economics*, 19(7), 539–550. <https://doi.org/10.1080/09603100801964370>
- Bodie, Z., Kane, A., & Marcus, A. J. (2018). *Investments (Eleventh edition)*. McGraw-Hill Education.
- Box, G. E. P. (o. J.). *Time Series Analysis*.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control (5th ed.)*. John Wiley & Sons.
- Campbell, J. Y., & Shiller, R. J. (1989). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *The Review of Financial Studies*, 1(3), 195–228.
- Carnot, N., Koen, V., & Tissot, B. (2005). *Economic Forecasting*. Palgrave Macmillan UK. <https://doi.org/10.1057/9780230005815>
- Castellani, M., & Santos, E. A. D. (o. J.). FORECASTING LONG-TERM GOVERNMENT BOND YIELDS: AN APPLICATION OF STATISTICAL AND AI MODELS
- Chaboud, A., Correia Golay, E., Cox, C., Fleming, M. J., Huh, Y., Keane, F. M., Lee, K., Schwarz, K., Vega, C., & Windover, C. (2025). All-to-All Trading in the U.S. Treasury Market. *Economic Policy Review (Federal Reserve Bank of New York)*, 31(2), 1–27. <https://doi.org/10.59576/epr.31.2.1-27>
- Clark, T. E., & McCracken, M. W. (1999). *Tests of equal forecast accuracy and encompassing for nested models (Working Paper)*. Federal Reserve Bank of Kansas City. <https://dx.doi.org/10.2139/ssrn.191028>
- McCracken, M. W. and Clark, T. E., *Advances in Forecast Evaluation (2011)*. Federal Reserve Bank of St. Louis Working Paper No. 2011-025B. <http://dx.doi.org/10.2139/ssrn.2038979>
- Cooray, A. (2002). The Fisher effect: A review of the literature. *Macquarie University, Department of Economics*.
- Cordeiro, W., Caldeira, J. F., & Moura, G. V. (o. J.). *Forecasting The Yield Curves Using Macroeconomics Expectations and Time-Varying Volatility*.

- Das, P. (2023). *Ökonometrie in Theorie und Praxis: Analyse von Querschnitt-, Zeitreihen- und Paneldaten mit Stata 15.1*. Springer Nature Singapore. <https://doi.org/10.1007/978-981-99-5940-2>
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/ N Portfolio Strategy? *Review of Financial Studies*, 22(5), 1915–1953. <https://doi.org/10.1093/rfs/hhm075>
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366), 427. <https://doi.org/10.2307/2286348>
- Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2), 337–364. <https://doi.org/10.1016/j.jeconom.2005.03.005>
- Diebold, F. X., & Rudebusch, G. D. (2013). *Yield curve modeling and forecasting: The dynamic Nelson-Siegel approach*. Princeton University Press.
- Diebold, F. X., Rudebusch, G. D., & Boragan Aruoba, S. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, 131(1–2), 309–338. <https://doi.org/10.1016/j.jeconom.2005.01.011>
- Duffee, G. R. (2012). Forecasting interest rates. *Economics Working Paper Archive 599, The Johns Hopkins University, Department of Economics*.
- Enders, W. (2015). *Applied econometric time series (4th ed.)*. John Wiley & Sons.
- Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251–276. <https://doi.org/10.2307/1913236>
- Estrella, A., & Hardouvelis, G. A. (1991). The term structure as a predictor of real economic activity. *Journal of Finance, American Finance Association, vol. 46(2), pages 555-576*. <https://doi.org/10.1111/j.1540-6261.1991.tb02674.x>
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417
- Federal Reserve Economic Data (FRED). (2025). *10-Year Treasury Constant Maturity Rate [Dataset]*. Federal Reserve Bank of St. Louis. <https://fred.stlouisfed.org/series/DGS10>
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669. <https://doi.org/10.1016/j.ejor.2017.11.054>
- Fisher, I. (1930). *The theory of interest*. Macmillan.
- Garbade, K. D. (2004). *Origins of the Federal Reserve Book-Entry System*. FRBNY Economic Policy Review.

- Garbade, K. D. (2012). *Birth of a market: The U.S. Treasury securities market from the Great War to the Great Depression*, The MIT Press.
- Gneiting, T., & Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>
- Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2(2), 111–120. [https://doi.org/10.1016/0304-4076\(74\)90034-7](https://doi.org/10.1016/0304-4076(74)90034-7)
- Granger, C. W. J. and Hyung, Namwon and Jeon, Y., Spurious Regressions with Stationary Series (1998). University of California at San Diego, Department of Economics Discussion Paper No. 98-25. <https://ssrn.com/abstract=145801>
- Greene, W. H. (2002). *Econometric analysis* (5th ed.). Prentice Hall.
- Guerard, J. B. (2013). *Introduction to Financial Forecasting in Investment Analysis*. Springer New York. <https://doi.org/10.1007/978-1-4614-5239-3>
- Guidolin, M., & Ionta, S. (2025). Forecasting Asset Returns Using Nelson–Siegel Factors Estimated from the US Yield Curve. *Econometrics*, 13(2), 17. <https://doi.org/10.3390/econometrics13020017>
- Gürkaynak, R. S., Sack, B., & Wright, J. H. (2006). *The U.S. Treasury yield curve: 1961 to the present* (Finance and Economics Discussion Series No. 2006-28). Federal Reserve Board.
- Hall, T., & Rasheed, K. (2025). A Survey of Machine Learning Methods for Time Series Prediction. *Applied Sciences*, 15(11), 5957. <https://doi.org/10.3390/app15115957>
- Hoesli, M., Lizieri, C., & MacGregor, B. (2007). The Inflation Hedging Characteristics of US and UK Investments: A Multi-Factor Error Correction Approach. *The Journal of Real Estate Finance and Economics*, 36(2), 183–206. <https://doi.org/10.1007/s11146-007-9062-6>
- Hoshmand, A. R. (2010). *Business forecasting: A practical approach* (Second edition). Routledge. <https://doi.org/10.4324/9780203874011>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. <https://otexts.com/fpp3/>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Investing.com. (2025). *US 10-Year T-Note Futures Historical Data* [Dataset]. <https://www.investing.com>
- Joshi, S. (o.J.). Dimensionality Reduction and State Space Systems: Forecasting the US Treasury Yields Using Frequentist and Bayesian VARs. SSRN Electronic Journal.

- Joshi, S. (o.J.). Forecasting the Leading Indicator of a Recession: The 10-Year Minus 3-Month Treasury Yield Spread. SSRN Electronic Journal.
- Koroleva, E., & Kopeykin, M. (2022). Understanding of Macro Factors That Affect Yield of Government Bonds. *Risks*, 10(8), 166. <https://doi.org/10.3390/risks10080166>
- Lemke, C., & Gabrys, B. (2010). Meta-learning for time series forecasting and forecast combination. *Neurocomputing*, 73(10–12), 2006–2016. <https://doi.org/10.1016/j.neucom.2009.09.020>
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.2307/2975974>
- Mishkin, F. S. (2019). *The economics of money, banking, and financial markets* (12th ed.). Pearson.
- Mönch, E. (2006). *Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach*. Humboldt University Berlin.
- Müller, C., Neyer, U., Wullweber, J., Schrooten, M., Schäfer, S., Sahr, A., Paetz, M., Hütten, M., Pigeaud, F., & Sylla, N. S. (o. J.). AUS POLITIK UND ZEITGESCHICHTE.
- Müller, T. (2013). *Finanzrisiken in der Assekuranz: Moderne Finanz- und Risikokonzepte in der Versicherungswirtschaft*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-8348-2307-6>
- Nelson, C. R., & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*, 60(4), 473–489. <http://www.jstor.org/stable/2352957>
- Neusser, K. (2025). *Time Series Econometrics*. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-88838-0>
- Palazzo, G., & Nobili, S. (2010). Explaining and Forecasting Bond Risk Premiums. *Financial Analysts Journal*, 66(4), 67–82. <https://doi.org/10.2469/faj.v66.n4.5>
- Paraskevopoulos, I., & Santos, A. (2002). The Stochastic Evolution of Financial Asset Prices. *Mathematics*, 13(12), 2002. <https://doi.org/10.3390/math13122002>
- Pesaran, M. H., & Timmermann, A. (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business & Economic Statistics*, 10(4), 461–465. <https://doi.org/10.1080/07350015.1992.10509922>
- Rocha, J., M. Viana, C., & Oliveira, S. (Hrsg.). (2024). *Time Series Analysis—Recent Advances, New Perspectives and Applications*. In *tech Open*. <https://doi.org/10.5772/intechopen.111223>

- Ryan, J. A., & Ulrich, J. M. (2008). xts: eXtensible Time Series (S. 0.14.1) [Dataset]. <https://doi.org/10.32614/CRAN.package.xts>
- Sahr, A. (Hrsg.). (2024). Geldpolitik im Umbruch. bpb, Bundeszentrale für politische Bildung.
- Said E. S., & Dickey, D. A. (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika*, 71(3), 599–607. <https://doi.org/10.2307/2336570>
- Sharpe, W.F. (1966) Mutual fund performance. *The Journal of Business*, 39, 119-138. <http://dx.doi.org/10.1086/294846>
- Sharpe, W.F. (1994) The Sharpe Ratio. *The Journal of Portfolio Management*, 21, 49-58. <http://dx.doi.org/10.3905/jpm.1994.409501>
- Shiller, R. (1980). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?(No. w0456; S. w0456). National Bureau of Economic Research. <https://doi.org/10.3386/w0456>
- Shumway, R. H., & Stoffer, D. S. (2018). *Time series analysis and its applications: With R examples* (4th ed.). Springer.
- Spiwoks, M., Bedke, N., & Hein, O. (2008). Forecasting the past: The case of US interest rate forecasts. *Financial Markets and Portfolio Management*, 22(4), 357–379. <https://doi.org/10.1007/s11408-008-0087-5>
- Stock, J. H. (with Watson, M. W.). (2020). *Introduction to econometrics* (Fourth edition, global edition). Pearson.
- Tan, Z., & Wu, Y. (2025). On Regime Switching Models. *Mathematics*, 13(7), 1128. <https://doi.org/10.3390/math13071128>
- Thome, H. (1994). Univariate Box/Jenkins models in time-series Analysis Univariate Box/Jenkins-Modelle in der Zeitreihenanalyse. *Historical Social Research*, 19, 577. <https://doi.org/10.12759>
- Tsay, R.S. (2010) *Analysis of Financial Time Series*. 3rd Edition, John Wiley & Sons, Hoboken. <https://doi.org/10.1002/9780470644560>
- Turlach, B. A., & Weingessel, A. (2019). *quadprog: Functions to solve quadratic programming problems* [R package]. <https://CRAN.R-project.org/package=quadprog>
- Venables, W.N. and Ripley, B.D. (2002) *Modern Applied Statistics with S*. Springer, New York, 271-300. <https://doi.org/10.1007/978-0-387-21706-2>
- Vo, N., & Ślepaczuk, R. (2022). Applying Hybrid ARIMA-SGARCH in Algorithmic Investment Strategies on S&P500 Index. *Entropy*, 24(2), 158. <https://doi.org/10.3390/e24020158>

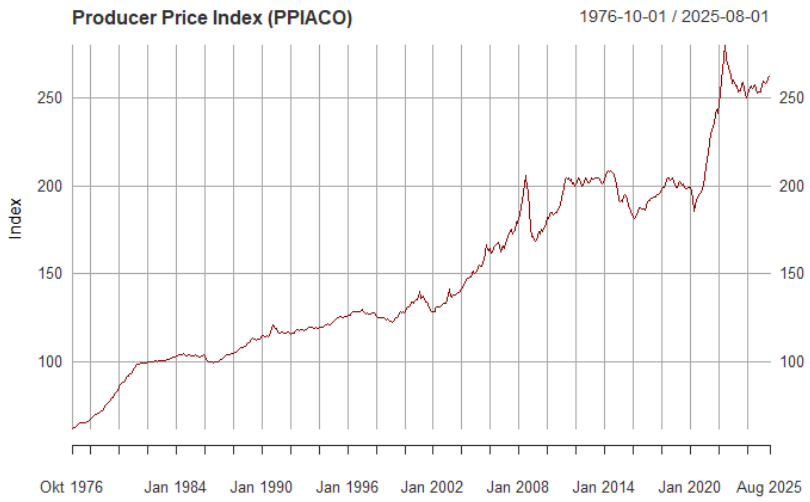
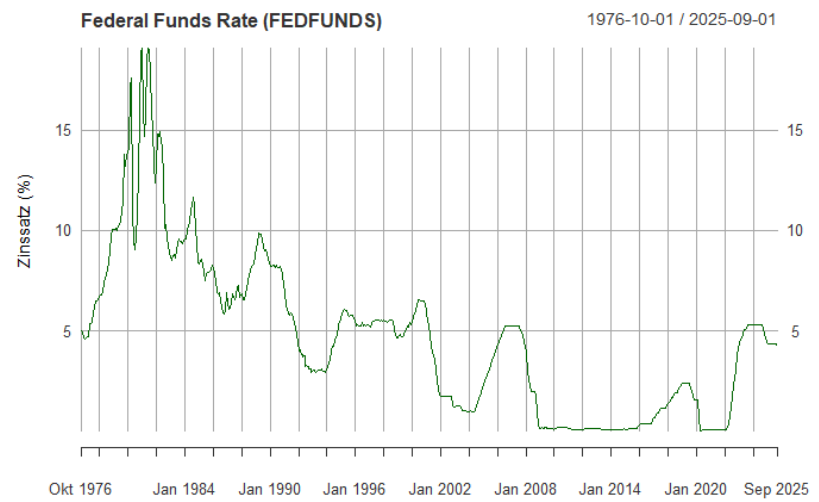
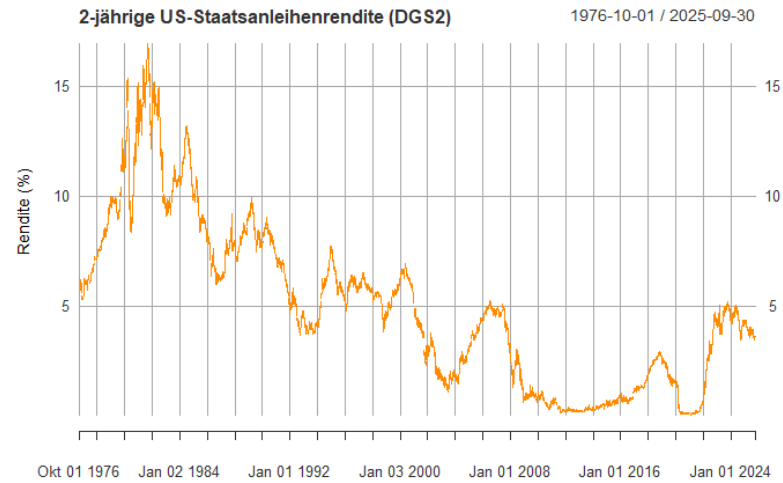
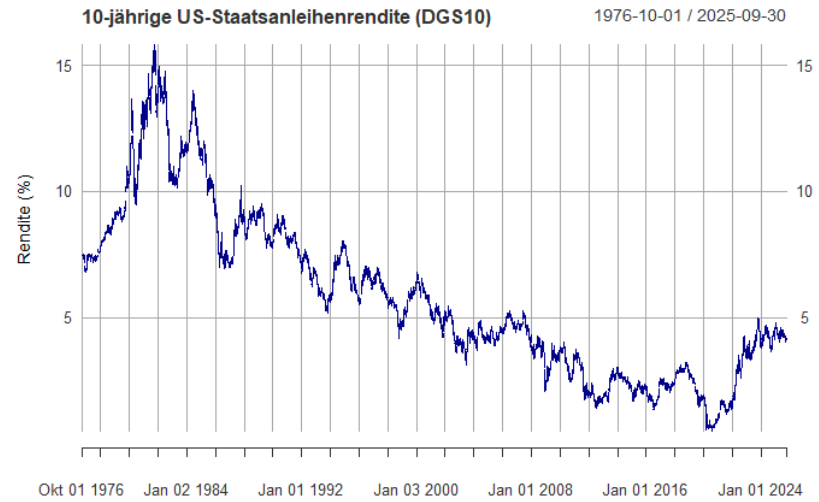
Wang, Y.-F., Wang, M. Y.-F., & Tu, L.-Y. (2025). An Evaluation of Machine Learning Models for Forecasting Short-Term U.S. Treasury Yields. *Applied Sciences*, 15(12), 6903. <https://doi.org/10.3390/app15126903>

Yahoo Finance. (2025). *S&P 500 Historical Data* [Dataset]. <https://finance.yahoo.com>

Zhang, Z. (2020). Research on Treasury Bond Futures Trading Strategy Based on ARIMA Model. *Finance and Market*, 5(3), 201. <https://doi.org/10.18686/fm.v5i3.2596>

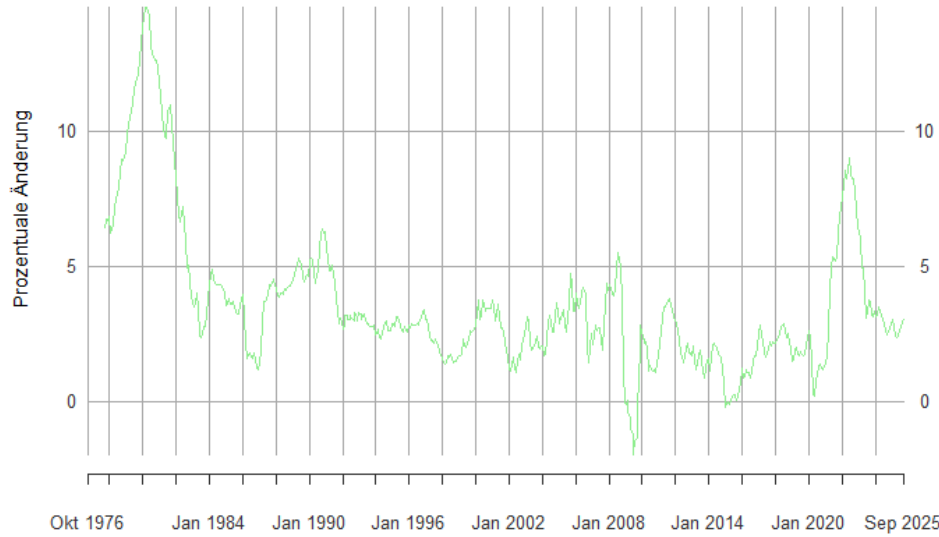
# Anhang

## Anhang A1: Ergebnisse des Augmented-Dickey-Fuller-Tests (ADF)



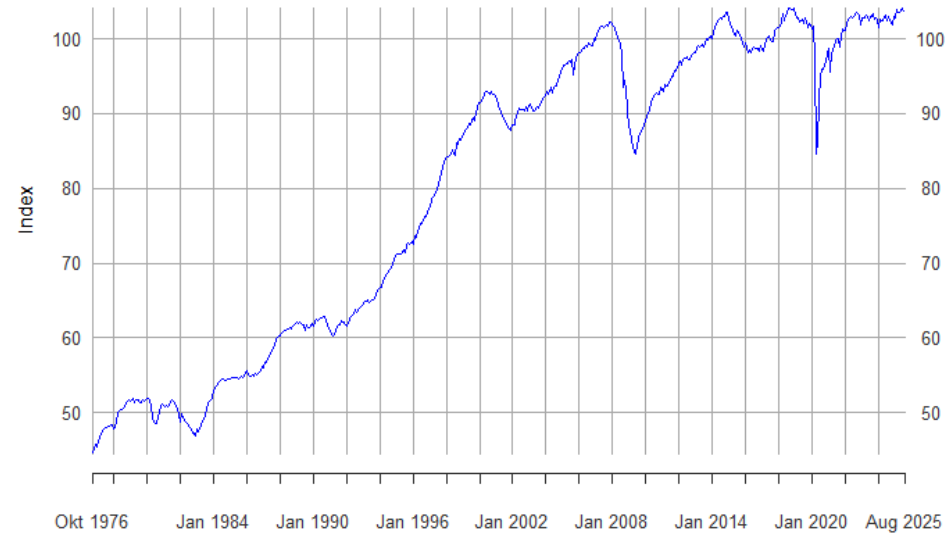
**Consumer Price Index (CPIAUCSL)**

1976-10-01 / 2025-09-01



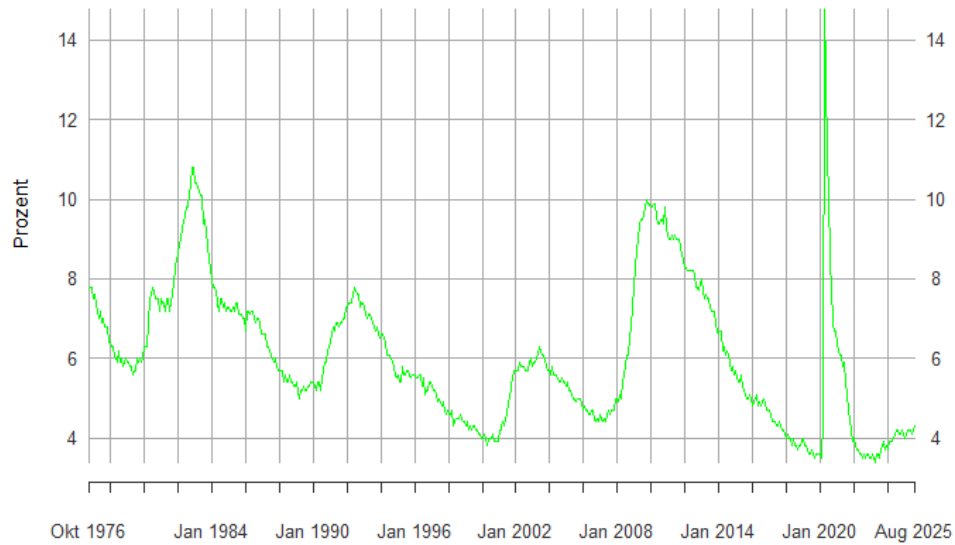
**Industrieproduktion (INDPRO)**

1976-10-01 / 2025-08-01



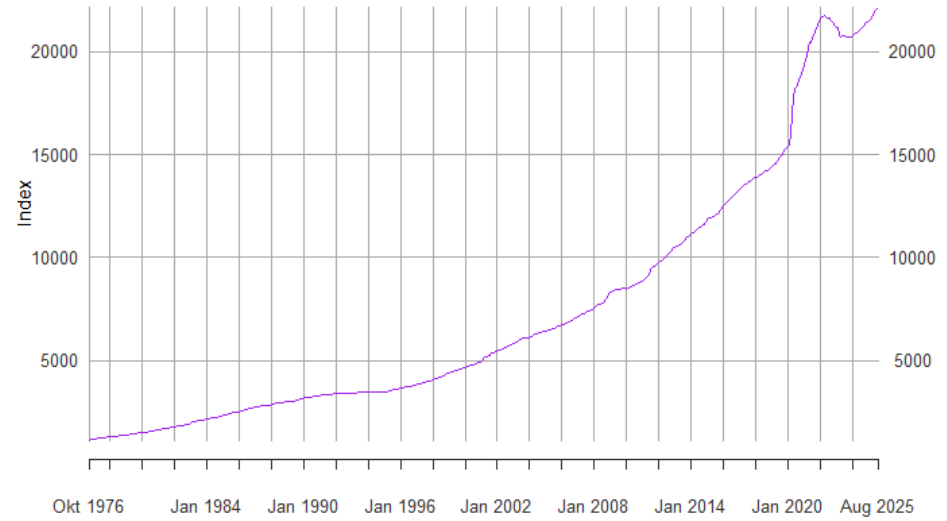
**Arbeitslosenquote (UNRATE)**

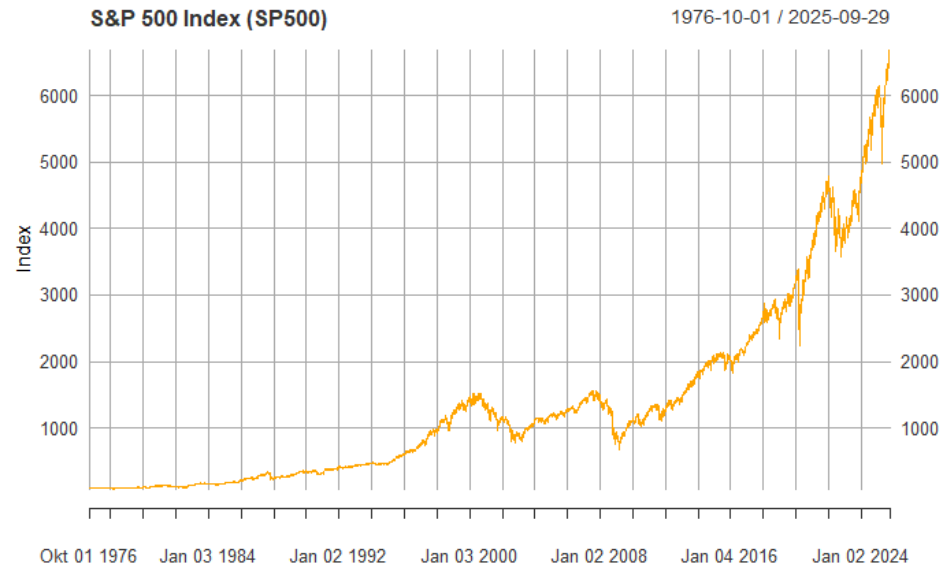
1976-10-01 / 2025-08-01



**M2 Geldmenge (M2SL)**

1976-10-01 / 2025-08-01





**Anhang A2: Ergebnisse des Augmented-Dickey-Fuller-Tests (ADF)**

**Anhang A2:** Ergebnisse des Augmented-Dickey-Fuller-Tests (ADF) nach Transformation der Variablen

<b>Variable</b>	<b>Teststatistik</b>	<b><i>p Wert</i></b>
d_DGS10	-7.499267	< 2.2e-16
d_DGS2	-7.415157	< 2.2e-16
d_FEDFUNDS	-6.319561	< 2.2e-16
CPI_g	-8.486575	< 2.2e-16
PPI_g	-4.145917	< 2.2e-16
INDPRO_g	-6.973035	< 2.2e-16
d_UNRATE	-4.716815	< 2.2e-16
M2_g	-7.692536	< 2.2e-16
SP500_r	-7.341657	< 2.2e-16

*Note.* Eigene Darstellung. Die Tabelle zeigt die für die Modellierung verwendeten makroökonomischen Variablen nach Anwendung der jeweiligen Transformationen zur Herstellung von Stationarität. Angegeben sind die ADF-Teststatistik ( $\tau$ ) und die zugehörigen p-Werte nach McKinnon. Logarithmierte Differenzen entsprechen Wachstumsraten, erste Differenzen Veränderungen in Prozentpunkten.

**Anhang A3: Modellselektion manuell und mit auto.arima()**

**Anhang A3: Modellselektion manuell und mit auto.arima() (R-Paket forecast)**

---

<b>Getestetes Modell</b>	<b>AIC</b>	<b><i>Log-Lokelihood</i></b>
arima(1,0,0)	331.56	-164.78
arima(0,0,1)	330.52	-164.26
arima(1,0,1)	332.08	-164.04
arima(2,0,0)	330.96	-163.48
arima(0,0,2)	331.7	-163.85
arima(2,0,1)	332.59	-163.29
arima(1,0,2)	333.1	-163.55
arima(2,0,2)	333.83	-162.92
auto.arima()	332.52	-164.26

---

*Note.* Manuelle Überprüfung und automatische Auswahl basierend auf dem Akaike Information Criterion (AIC) unter Verwendung der Funktion auto.arima() aus dem R-Paket forecast (Hyndman & Athanasopoulos, 2021, ch. 8.5). Das MA(1)-Modell wurde aufgrund des niedrigsten AIC-Werts ausgewählt.

**Anhang A4:** Ergebnisse der ARIMA(0,0,1)-Schätzung (DGS10-Rendite)

```
Series: train_data$d_DGS10_centered
ARIMA(0,0,1) with zero mean
```

```
Coefficients:
```

```
      ma1
      0.1578
s.e. 0.0554
```

```
sigma^2 = 0.153: log likelihood = -164.26
AIC=332.52  AICc=332.56  BIC=340.2
```

```
Training set error measures:
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-0.001428037	0.3905972	0.2882545	87.79055	133.7653	0.7450351	-0.007036206

Eigene Berechnung in R mit dem Paket forecast (Hyndman & Athanasopoulos, 2021, ch. 8.5). Der Screenshot zeigt den vollständigen R-Output der Modellschätzung für das ARIMA(0,0,1)-Modell der 10-jährigen US-Treasury-Rendite. Der positive MA(1)-Koeffizient ( $\theta_1 = 0.158$ ,  $p < 0.001$ ) weist auf eine leichte kurzfristige Persistenz hin.

## Anhang A5: Signifikanznachweis der fundamentalen Modelle (Modell 1–3)

### A5.1 Regressionsausgabe des Modell 1

Call:

```
lm(formula = d_DGS10 ~ d_DGS10_lag1 + d_DGS2_lag1 + d_FEDFUNDS_lag1 +
    d_UNRATE_lag1 + CPI_g_lag1 + INDPRO_g_lag1 + M2_g_lag1 +
    SP500_r_lag1 + PPI_g_lag1, data = train_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.65210	-0.19593	0.00286	0.18920	1.23067

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.105533	0.045213	-2.334	0.02018 *
d_DGS10_lag1	0.118690	0.129049	0.920	0.35838
d_DGS2_lag1	-0.003239	0.095956	-0.034	0.97310
d_FEDFUNDS_lag1	-0.005545	0.036996	-0.150	0.88094
d_UNRATE_lag1	0.048327	0.140932	0.343	0.73188
CPI_g_lag1	0.034216	0.086967	0.393	0.69425
INDPRO_g_lag1	0.084823	0.038101	2.226	0.02667 *
M2_g_lag1	0.082220	0.060938	1.349	0.17818
SP500_r_lag1	0.016372	0.004962	3.299	0.00108 **
PPI_g_lag1	0.057128	0.037058	1.542	0.12412

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3851 on 333 degrees of freedom  
Multiple R-squared: 0.07705, Adjusted R-squared: 0.0521  
F-statistic: 3.089 on 9 and 333 DF, p-value: 0.001411

### A5.2 Regressionsausgabe des Modell 2

Call:

```
lm(formula = d_DGS10 ~ d_DGS2_lag1 + d_FEDFUNDS_lag1 + M2_g_lag1 +
    SP500_r_lag1 + PPI_g_lag1, data = train_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.73449	-0.19381	0.00119	0.19185	1.34475

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.075296	0.037669	-1.999	0.0464 *
d_DGS2_lag1	0.091776	0.041650	2.204	0.0282 *
d_FEDFUNDS_lag1	0.007896	0.033757	0.234	0.8152
M2_g_lag1	0.079172	0.060162	1.316	0.1891
SP500_r_lag1	0.014749	0.004880	3.023	0.0027 **
PPI_g_lag1	0.066346	0.029970	2.214	0.0275 *

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3864 on 337 degrees of freedom  
Multiple R-squared: 0.05968, Adjusted R-squared: 0.04573  
F-statistic: 4.278 on 5 and 337 DF, p-value: 0.0008726

### A5.3 Regressionsausgabe des Modell 3

```
call:
lm(formula = d_DGS10 ~ d_DGS10_lag1 + d_DGS10_lag2 + d_DGS10_lag3 +
  d_DGS2_lag1 + d_DGS2_lag2 + d_DGS2_lag3 + d_FEDFUNDS_lag1 +
  d_FEDFUNDS_lag2 + d_FEDFUNDS_lag3 + d_UNRATE_lag1 + d_UNRATE_lag2 +
  d_UNRATE_lag3 + CPI_g_lag1 + CPI_g_lag2 + CPI_g_lag3 + INDPRO_g_lag1 +
  INDPRO_g_lag2 + INDPRO_g_lag3 + M2_g_lag1 + M2_g_lag2 + M2_g_lag3 +
  SP500_r_lag1 + SP500_r_lag2 + SP500_r_lag3 + PPI_g_lag1 +
  PPI_g_lag2 + PPI_g_lag3, data = train_data)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-1.48938 -0.18686 -0.00027  0.20366  1.20374
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.111016   0.055797  -1.990 0.047494 *
d_DGS10_lag1  0.152136   0.132816   1.145 0.252884
d_DGS10_lag2  0.112547   0.138254   0.814 0.416224
d_DGS10_lag3  0.229303   0.139966   1.638 0.102363
d_DGS2_lag1  -0.056406   0.101633  -0.555 0.579289
d_DGS2_lag2  -0.205717   0.112039  -1.836 0.067283 .
d_DGS2_lag3  -0.262927   0.112470  -2.338 0.020026 *
d_FEDFUNDS_lag1  0.056933   0.049404   1.152 0.250033
d_FEDFUNDS_lag2  0.095113   0.047399   2.007 0.045641 *
d_FEDFUNDS_lag3  0.064690   0.040954   1.580 0.115205
d_UNRATE_lag1  -0.053290   0.151092  -0.353 0.724548
d_UNRATE_lag2  -0.012318   0.150438  -0.082 0.934792
d_UNRATE_lag3  -0.011258   0.147643  -0.076 0.939269
CPI_g_lag1     0.032759   0.113575   0.288 0.773202
CPI_g_lag2    -0.057937   0.119492  -0.485 0.628111
CPI_g_lag3     0.102421   0.109183   0.938 0.348926
INDPRO_g_lag1  0.118930   0.039652   2.999 0.002922 **
INDPRO_g_lag2  0.004548   0.040105   0.113 0.909791
INDPRO_g_lag3 -0.082864   0.040192  -2.062 0.040054 *
M2_g_lag1     -0.024111   0.078812  -0.306 0.759860
M2_g_lag2     0.144467   0.083674   1.727 0.085228 .
M2_g_lag3    -0.042976   0.075203  -0.571 0.568086
SP500_r_lag1  0.017672   0.005051   3.499 0.000535 ***
SP500_r_lag2  0.006793   0.005087   1.335 0.182769
SP500_r_lag3  0.005398   0.005103   1.058 0.290969
PPI_g_lag1    0.077017   0.038036   2.025 0.043726 *
PPI_g_lag2   -0.061342   0.038638  -1.588 0.113378
PPI_g_lag3    0.017726   0.039302   0.451 0.652279
---
```

```
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.3795 on 315 degrees of freedom
Multiple R-squared:  0.1524,    Adjusted R-squared:  0.07977
F-statistic: 2.098 on 27 and 315 DF,  p-value: 0.001472
```

Eigene Berechnung in R. Der Anhang zeigt die vollständigen Regressionsausgaben (lm-Funktion) der fundamentalen Modelle 1 bis 4, wie in Tabelle 4 beschrieben. Dargestellt sind die geschätzten Koeffizienten (Estimate), Standardfehler (Std. Error), t-Werte und zugehörigen p-Werte der jeweiligen Spezifikationen. Die Signifikanzcodes am Tabellenende entsprechen der R-Konvention:

" =  $p < 0.001$ , " =  $p < 0.01$ , " =  $p < 0.05$ , '.' =  $p < 0.1$ .

**Anhang A6:** Engle-Granger-Kointegrationstest zwischen DGS10 und DGS2

## Augmented Dickey-Fuller Test

```
data: as.numeric(residuals_ecm_xts)
Dickey-Fuller = -3.4776, Lag order = 8, p-value = 0.04446
alternative hypothesis: stationary
```

Eigene Berechnung in R. Der Screenshot zeigt das Ergebnis des zweiten Schritts des Engle-Granger-Verfahrens zur Überprüfung einer langfristigen Gleichgewichtsbeziehung zwischen den Renditen 10-jähriger (DGS10) und 2-jähriger (DGS2) US-Staatsanleihen. Der Augmented-Dickey-Fuller-Test (ADF) auf die Residuen ergibt eine Teststatistik von  $-3.4776$  und einen p-Wert von  $0.044$ , womit die Nullhypothese einer Einheitswurzel auf dem 5 %-Niveau verworfen werden kann. Damit sind die Residuen stationär und eine Kointegrationsbeziehung bestätigt.

**Anhang A7: Regressionsausgabe des ECM<sub>2</sub> (Error-Correction-Ansatz)**

```

call:
lm(formula = d_DGS10 ~ ECM_Residuals_lag1 + d_DGS10_lag1 + d_DGS2_lag1 +
    SP500_r_lag1, data = train_data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.72276 -0.20492 -0.01273  0.20386  1.37041

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.013510   0.021949  -0.616  0.53863
ECM_Residuals_lag1 -0.021436   0.031066  -0.690  0.49066
d_DGS10_lag1     0.173897   0.121296   1.434  0.15259
d_DGS2_lag1     -0.004445   0.088961  -0.050  0.96018
SP500_r_lag1     0.015147   0.004937   3.068  0.00233 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3883 on 338 degrees of freedom
Multiple R-squared:  0.04755, Adjusted R-squared:  0.03628
F-statistic: 4.218 on 4 and 338 DF, p-value: 0.002401

```

Eigene Darstellung. Die Tabelle zeigt die Schätzergebnisse des Error-Correction-Modells (ECM<sub>2</sub>) zur Erklärung der Veränderung der zehnjährigen US-Treasury-Rendite ( $\Delta DGS_{10}$ ). Der Fehlerkorrekturterm ( $ECM\_Residuals_{t-1}$ ) weist mit  $\gamma = -0,021$  ( $p = 0,491$ ) das erwartete negative Vorzeichen auf, ist jedoch statistisch insignifikant und deutet auf eine langsame Rückkehr zum langfristigen Gleichgewicht hin. Kurzfristig zeigen die Renditeveränderungen des S&P 500 ( $\beta = 0,015$ ;  $p = 0,002$ ) einen signifikant positiven Einfluss, während die verzögerten Veränderungen der zehnjährigen ( $\Delta DGS_{10,t-1}$ ) und zweijährigen Rendite ( $\Delta DGS_{2,t-1}$ ) keine signifikanten Effekte aufweisen. Das Modell erklärt rund 4,8 % der Varianz ( $R^2 = 0,0475$ ) und ist insgesamt signifikant ( $F(4, 338) = 4,218$ ;  $p = 0,0024$ ).

**Anhang A8: Performancekennzahlen aller modellbasierten Handelsstrategien****Anhang A8: Performancekennzahlen aller modellbasierten Handelsstrategien**

Modell	<i>Cumulative Return</i>	Ann. Return	Ann. Volatility	Sharpe Ratio	Max DD
ECM_2	1.0773	0.0391	0.0572	0.6831	0.0835
ARIMAX	0.8201	0.0319	0.0575	0.5541	0.0962
Modell_2	0.6524	0.0267	0.0578	0.4617	0.1676
ECM_3	0.5918	0.0247	0.0578	0.4264	0.1006
Manual Final	0.5793	0.0242	0.0579	0.4189	0.1676
Modell_4	0.5650	0.0237	0.0579	0.4104	0.1192
Modell_1	0.5333	0.0226	0.0579	0.3911	0.2031
ARMA Rek. Opt.	0.5215	0.0222	0.0579	0.3839	0.1714
ARIMAX Rol. 20J	0.4409	0.0193	0.0580	0.3331	0.1240
ECM_5	0.3354	0.0153	0.0581	0.2628	0.2102
Modell_3	0.3309	0.0151	0.0581	0.2597	0.2103
ARIMAX Rol. 15J	0.3185	0.0146	0.0581	0.2510	0.1098
Stepwise	0.2615	0.0122	0.0582	0.2105	0.2786
ECM_4	0.2615	0.0122	0.0582	0.2105	0.2786
ARIMAX Rol. 2J	0.2055	0.0098	0.0582	0.1690	0.1749
ECM_1	0.1814	0.0088	0.0582	0.1506	0.1584
ARIMAX Rol. 10J	0.1410	0.0069	0.0583	0.1191	0.1176
ARMA Rol. Fest 20J	0.0772	0.0039	0.0583	0.0670	0.1633
MA(1) Rol. 10J	0.0529	0.0027	0.0583	0.0464	0.2058
Buy and Hold Benchmark	0.0233	0.0012	0.0583	0.0207	0.2448

---

MA(1) Rol. 20J	-0.0708	-0.0038	0.0583	-0.0658	0.1875
MA(1) Rol. 15J	-0.0778	-0.0042	0.0583	-0.0726	0.1936
ARIMAX Rol. 5J	-0.1056	-0.0058	0.0583	-0.1000	0.2483
MA(1) Rekursiv	-0.1127	-0.0062	0.0583	-0.1071	0.2224
Naive	-0.1959	-0.0114	0.0583	-0.1951	0.2777
MA(1) Rol. 5J	-0.2012	-0.0117	0.0583	-0.2009	0.2995
MA(1) Rol. 2J	-0.2236	-0.0132	0.0582	-0.2262	0.2852
ARMA Rol. Opt. 20J	-0.3044	-0.0188	0.0581	-0.3242	0.3744

---

*Note.* Eigene Darstellung. Die Tabelle zeigt die Ergebnisse der modellbasierten Long/Short-Handelsstrategien auf Basis der Prognosen der univariaten Modelle. Dargestellt sind der kumulierte Ertrag (Cumulative Return), die annualisierte Rendite (Ann. Return), die annualisierte Volatilität (Ann. Volatility), die Sharpe Ratio sowie der maximale Drawdown (Max DD). Höhere Sharpe Ratios und geringere Drawdowns weisen auf ein günstigeres Risiko-Rendite-Profil hin, während die annualisierte Rendite den durchschnittlichen jährlichen Ertrag der jeweiligen Strategie beschreibt.